第一章　凸集

凸集是凸分析（convex analysis）中最重要最基本的概念之一．本章首先研究凸集的概念与性质，然后研究一些特殊的凸集：凸包、凸多面体、凸代数体、凸体、凸锥，最后介绍$\mathscr{R}^n$中的几个经典定理与凸集的相对内部．

在本书中用$\mathscr{R}$表示全部实数集合，V表示$\mathscr{R}$上的线性空间，$\mathscr{R}^n$表示n维欧氏空间．为方便起见，假定V中包含的点多于一个．在$\mathscr{R}^n$中向量范数$\|\boldsymbol{x}\|$为欧氏范数．

设$\boldsymbol{x}$，$\boldsymbol{y}\in V$，则集合$\{\lambda\boldsymbol{x}+(1-\lambda)\boldsymbol{y}\mid 0\leqslant\lambda\leqslant 1\}=\{\boldsymbol{y}+\lambda(\boldsymbol{x}-\boldsymbol{y})\mid 0\leqslant\lambda\leqslant 1\}$表示以$\boldsymbol{x}$，$\boldsymbol{y}$为端点的线段，记为$[\boldsymbol{x},\ \boldsymbol{y}]$．与此类似，$(\boldsymbol{x},\ \boldsymbol{y})=\{\lambda\boldsymbol{x}+(1-\lambda)\boldsymbol{y}\mid 0<\lambda<1\}$，$(\boldsymbol{x},\ \boldsymbol{y}]=\{\lambda\boldsymbol{x}+(1-\lambda)\boldsymbol{y}\mid 0\leqslant\lambda<1\}$，$[\boldsymbol{x},\ \boldsymbol{y})=\{\lambda\boldsymbol{x}+(1-\lambda)\boldsymbol{y}\mid 0<\lambda\leqslant 1\}$．

若$\boldsymbol{x}=\boldsymbol{y}$，则$[\boldsymbol{x},\ \boldsymbol{y}]=[\boldsymbol{x},\ \boldsymbol{y})=(\boldsymbol{x},\ \boldsymbol{y}]=(\boldsymbol{x},\ \boldsymbol{y})=\{\boldsymbol{x}\}$．

若x，$y\in\mathscr{R}$，且$x\neq y$，则记号$[x,\ y]$，$[x,\ y)$，$(x,\ y]$以及$(x,\ y)$的含义与数学分析中的规定相同．

§1.1　凸集·凸包·凸多面体

1.1.1　凸集

定义 1.1.1　若$\boldsymbol{x}^{(i)}\in V$，$\lambda_i\geqslant 0$，$i=1,\ \cdots,\ k$，且$\sum\limits_{i=1}^{k}\lambda_i=1$，则称点$\boldsymbol{x}=\sum\limits_{i=1}^{k}\lambda_i\boldsymbol{x}^{(i)}$为$\boldsymbol{x}^{(1)}$，$\cdots$，$\boldsymbol{x}^{(k)}$的**凸组合**．

定义 1.1.2　设$A\subset V$，若$\forall\boldsymbol{x}$，$\boldsymbol{y}\in A$均有$(\boldsymbol{x},\ \boldsymbol{y})=\{\lambda\boldsymbol{x}$

$+(1-\lambda)\boldsymbol{y} | 0<\lambda<1\} \subset A$，则称$A$为**凸集**（convex set）.

可以验证，以下各例中的集合都是凸集.

例1 空集Φ，单元素集$\{\boldsymbol{a}\}$，整个线性空间V（这些都是平凡的凸集）.

例2 $S=\{\boldsymbol{x} | A\boldsymbol{x}=\boldsymbol{b},\ \boldsymbol{x}\geqslant \boldsymbol{0}\}$，其中$A$为$m\times n$矩阵，$\boldsymbol{b}\in \mathscr{R}^m$，$\boldsymbol{x}\in \mathscr{R}^n$.

例3 超球 $B(\boldsymbol{a};\ r)=\{\boldsymbol{x} | \boldsymbol{x}\in \mathscr{R}^n,\ \|\boldsymbol{x}-\boldsymbol{a}\|<r\}$.

例4 $A=\{$所有正系数单元多项式$\}$（设V为$\mathscr{R}$上的所有单元多项式构成的线性空间）.

凸集具有以下性质：

性质1 设$C_1,\ C_2,\ \cdots,\ C_k\subset V$均为凸集，$\alpha_i\in \mathscr{R}$，$i=1,\ \cdots,\ k$，则$\sum\limits_{i=1}^{k} \alpha_i C_i$ 为凸集.

$$\left(\sum_{i=1}^{k} \alpha_i C_i=\left\{\boldsymbol{x} \middle| \boldsymbol{x}=\sum_{i=1}^{k} \alpha_i \boldsymbol{x}^{(i)},\ \boldsymbol{x}^{(i)}\in C_i\right\}\right)$$

性质2 若C_j为凸集，$j\in J$，则它们的交$\bigcap\limits_{j\in J} C_j$也是凸集，其中$J$为任一指标集.

（证明留给读者）

性质3 设$A\subset V$，则A是**凸**集等价于

$$(\alpha+\beta)A=\alpha A+\beta A,\quad \forall\, \alpha,\ \beta\geqslant 0. \tag{1.1-1}$$

证 1° 设A是凸集，再设α，β是不同时为零的任意两个非负实数.

设$\boldsymbol{z}\in(\alpha+\beta)A$，即存在$\boldsymbol{x}\in A$，使$\boldsymbol{z}=(\alpha+\beta)\boldsymbol{x}=\alpha\boldsymbol{x}+\beta\boldsymbol{x}$. 因为$\alpha\boldsymbol{x}\in\alpha A$，$\beta\boldsymbol{x}\in\beta A$，所以$\boldsymbol{z}\in\alpha A+\beta A$，因而

$$(\alpha+\beta)A\subset\alpha A+\beta A.$$

设$\boldsymbol{z}\in\alpha A+\beta A$，即存在$\boldsymbol{x}^{(1)}$，$\boldsymbol{x}^{(2)}\in A$，使

$$\boldsymbol{z}=\alpha\boldsymbol{x}^{(1)}+\beta\boldsymbol{x}^{(2)}$$

$$=(\alpha+\beta)\left(\frac{\alpha}{\alpha+\beta}\boldsymbol{x}^{(1)}+\frac{\beta}{\alpha+\beta}\boldsymbol{x}^{(2)}\right).$$

由于A是凸集，$\boldsymbol{x}^{(1)}$，$\boldsymbol{x}^{(2)}\in A$，所以$\boldsymbol{x}^{(1)}$，$\boldsymbol{x}^{(2)}$的凸组合属于$A$，即

$$\frac{\alpha}{\alpha+\beta}\boldsymbol{x}^{(1)}+\frac{\beta}{\alpha+\beta}\boldsymbol{x}^{(2)}\in A,$$

因而　$\boldsymbol{z}\in(\alpha+\beta)A$，于是$\alpha A+\beta A\subset(\alpha+\beta)A$。

综上所述得式(1.1-1)（当$\alpha=\beta=0$时，式(1.1-1)显然也成立）.

2° 假设式(1.1-1)成立，再设$\boldsymbol{x}^{(1)}$，$\boldsymbol{x}^{(2)}\in A$，$\lambda\in(0,1)$。任取$\alpha>0$，令$\beta=\left(\frac{1}{\lambda}-1\right)\alpha$. 由此可得$\lambda=\frac{\alpha}{\alpha+\beta}$，$1-\lambda=\frac{\beta}{\alpha+\beta}$.

由式(1.1-1)得　$\alpha\boldsymbol{x}^{(1)}+\beta\boldsymbol{x}^{(2)}\in(\alpha+\beta)A$，即存在$\boldsymbol{x}\in A$使$\alpha\boldsymbol{x}^{(1)}+\beta\boldsymbol{x}^{(2)}=(\alpha+\beta)\boldsymbol{x}$. 由此得

$$\lambda\boldsymbol{x}^{(1)}+(1-\lambda)\boldsymbol{x}^{(2)}=\frac{\alpha}{\alpha+\beta}\boldsymbol{x}^{(1)}+\frac{\beta}{\alpha+\beta}\boldsymbol{x}^{(2)}=\boldsymbol{x}.$$

因为$\boldsymbol{x}\in A$，所以$\lambda\boldsymbol{x}^{(1)}+(1-\lambda)\boldsymbol{x}^{(2)}\in A$，因而$A$是凸集.

性质4　设W是线性空间，T是由V到W的线性映射.

1° 若$A\subset V$是凸集，则$TA\subset W$是凸集；

2° 若$B\subset W$是凸集，则$T^{-1}B\subset V$是凸集.

证　1° 设A是凸集，$\lambda\in(0,1)$. 若$\boldsymbol{y}^{(1)}$，$\boldsymbol{y}^{(2)}\in TA$，则存在$\boldsymbol{x}^{(1)}$，$\boldsymbol{x}^{(2)}\in A$，使

$$\boldsymbol{y}^{(i)}=T\boldsymbol{x}^{(i)},\quad i=1,\ 2.$$

因而

$$\begin{aligned}\lambda\boldsymbol{y}^{(1)}+(1-\lambda)\boldsymbol{y}^{(2)}&=\lambda T\boldsymbol{x}^{(1)}+(1-\lambda)T\boldsymbol{x}^{(2)}\\&=T(\lambda\boldsymbol{x}^{(1)}+(1-\lambda)\boldsymbol{x}^{(2)}).\end{aligned}$$

由A的凸性可得 $\lambda\boldsymbol{x}^{(1)}+(1-\lambda)\boldsymbol{x}^{(2)}\in A$. 于是 $\lambda\boldsymbol{y}^{(1)}+(1-\lambda)\boldsymbol{y}^{(2)}$

$\in TA$. 所以TA为凸集.

2° 设$B\subset W$是凸集，$\lambda\in(0,1)$. 若$x^{(1)}$, $x^{(2)}\in T^{-1}B$，则存在$y^{(1)}$, $y^{(2)}\in B$，使$T^{-1}y^{(i)}=x^{(i)}$, $i=1, 2$, 因而

$$\lambda x^{(1)}+(1-\lambda)x^{(2)}=\lambda T^{-1}y^{(1)}+(1-\lambda)T^{-1}y^{(2)}$$
$$=T^{-1}(\lambda y^{(1)}+(1-\lambda)y^{(2)}).$$

由B的凸性可得$\lambda y^{(1)}+(1-\lambda)y^{(2)}\in B$. 于是 $\lambda x^{(1)}+(1-\lambda)x^{(2)}\in T^{-1}B$. 所以$T^{-1}B$为凸集.

1.1.2 凸包

定义 1.1.3 设$A\subset V$. V中所有包含A的凸子集的交称为A的**凸包**(convex hull)，记为co(A)（见图1-1）.

显然，凸包co(A)是包含A的最小凸集，凸包具有以下重要性质：

定理 1.1.1 若$A\subset V$, $x\in \mathrm{co}(A)$，则

$$\mathrm{co}(A\cup\{x\})=\mathrm{co}(A).$$

证 因为 $A\subset\mathrm{co}(A)$, $x\in\mathrm{co}(A)$，所以

$$A\cup\{x\}\subset\mathrm{co}(A),$$

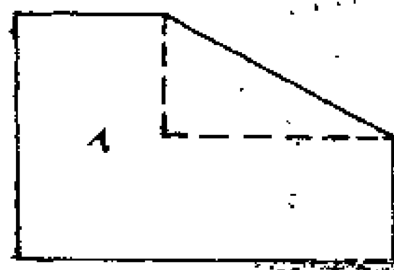


图1-1

因而

$$\mathrm{co}(A\cup\{x\})\subset\mathrm{co}(A).$$

显然还有

$$\mathrm{co}(A)\subset\mathrm{co}(A\cup\{x\}),$$

于是

$$\mathrm{co}(A\cup\{x\})=\mathrm{co}(A).$$

定理 1.1.2 设$A\subset V$，则A的凸包co(A)等于A中元素的有限凸组合的全体.

证 设B为A中元素的有限凸组合的全体.

1° 设$x^{(1)}, \cdots, x^{(k)}\in A$，因而也属于co($A$). 于是

$$\sum_{i=1}^{k}\lambda_i x^{(i)}\in\mathrm{co}(A),\ \forall\lambda_i\geqslant 0,\quad \sum_{i=1}^{k}\lambda_i=1.$$

（见本章习题2）因而$B\subset \text{co}(A)$.

2° 设$\boldsymbol{x}$，$\boldsymbol{y}\in B$，即存在$\boldsymbol{x}^{(i)}$，$\boldsymbol{y}^{(j)}\in A$，$\lambda_i\geqslant 0$，$\mu_j\geqslant 0$，$\sum_{i=1}^{k}\lambda_i=1$，$\sum_{j=1}^{m}\mu_j=1$使

$$\boldsymbol{x}=\sum_{i=1}^{k}\lambda_i\boldsymbol{x}^{(i)},\ \boldsymbol{y}=\sum_{j=1}^{m}\mu_j\boldsymbol{y}^{(j)}.$$

$\boldsymbol{x}$，$\boldsymbol{y}$的凸组合

$$\lambda\boldsymbol{x}+(1-\lambda)\boldsymbol{y}=\sum_{i=1}^{k}\lambda\lambda_i\boldsymbol{x}^{(i)}+\sum_{j=1}^{m}(1-\lambda)\mu_j\boldsymbol{y}^{(j)},\ \lambda\in(0,\ 1)$$

而 $\lambda\lambda_i\geqslant 0$，$(1-\lambda)\mu_j\geqslant 0$，且

$$\sum_{i=1}^{k}\lambda\lambda_i+\sum_{j=1}^{m}(1-\lambda)\mu_j=\lambda+(1-\lambda)=1,$$

所以 $\lambda\boldsymbol{x}+(1-\lambda)\boldsymbol{y}\in B$，因而$B$为凸集．而$B\supset A$，所以$B\supset \text{co}(A)$。

综上所述得 $B=\text{co}(A)$。

定理 1.1.3 设A，$B\subset V$是凸集，则

$$\text{co}(A\cup B)=\bigcup_{0\leqslant\lambda\leqslant 1}[\lambda A+(1-\lambda)B].$$

证 1° 设$\boldsymbol{z}\in\bigcup_{0\leqslant\lambda\leqslant 1}[\lambda A+(1-\lambda)B]$，因而存在$\lambda_1\in[0,1]$，使$\boldsymbol{z}\in\lambda_1 A+(1-\lambda_1)B$. 于是存在$\boldsymbol{x}^{(1)}\in A$，$\boldsymbol{y}^{(1)}\in B$，使

$$\boldsymbol{z}=\lambda_1\boldsymbol{x}^{(1)}+(1-\lambda_1)\boldsymbol{y}^{(1)}.$$

又因为 $\boldsymbol{x}^{(1)}$，$\boldsymbol{y}^{(1)}\in \text{co}(A\cup B)$，$\lambda_1\in[0,\ 1]$，所以$\boldsymbol{z}\in \text{co}(A\cup B)$，因而

$$\bigcup_{0\leqslant\lambda\leqslant 1}[\lambda A+(1-\lambda)B]\subset \text{co}(A\cup B).$$

2° 设$\boldsymbol{z}\in \text{co}(A\cup B)$. 依定理 1.1.2，存在$\boldsymbol{z}^{(i)}\in A\cup B$，$\lambda_i\geqslant$

0， $\sum_{i=1}^{m} \lambda_i = 1$，使 $z = \sum_{i=1}^{m} \lambda_i z^{(i)}$。

若 $z^{(i)} \in A$，$i = 1, \cdots, m$，则 $z \in A$（见本章习题2）。因而 $z \in \bigcup_{0 \leqslant \lambda \leqslant 1} [\lambda A + (1-\lambda) B]$。若 $z^{(i)} \in B$，$i = 1, \cdots, m$，也有同样结论。

设 $z^{(i)} \in A$，$i = 1, \cdots, k$；$z^{(i)} \in B$，$i = k+1, \cdots, m$。令

$$\lambda = \lambda_1 + \cdots + \lambda_k, \quad 0 < \lambda < 1$$

于是

$$z = \sum_{i=1}^{k} \lambda_i z^{i} + \sum_{i=k+1}^{m} \lambda_i z^{i}$$

$$= \lambda \sum_{i=1}^{k} \frac{\lambda_i}{\lambda} z^{i} + (1-\lambda) \sum_{i=k+1}^{m} \frac{\lambda_i}{1-\lambda} z^{(i)}.$$

因为 A 为凸集，$z^{i} \in A$，$i = 1, \cdots, k$，所以存在 $x \in A$，使

$$x = \sum_{i=1}^{k} \frac{\lambda_i}{\lambda} z^{i}.$$

同样，存在 $y \in B$，使

$$y = \sum_{i=k+1}^{m} \frac{\lambda_i}{1-\lambda} z^{i}.$$

于是 $z = \lambda x + (1-\lambda) y$，因而 $z \in \lambda A + (1-\lambda) B$，由此可得

$$\mathrm{co}(A \cup B) \subset \bigcup_{0 \leqslant \lambda \leqslant 1} [\lambda A + (1-\lambda) B].$$

综合1°与2°，命题得证。

1.1.3 凸多面体

V 中有限点集 $A = \{x^{1}, \cdots, x^{m}\}$ 的凸包记为 $\mathrm{co}(x^{(1)}, \cdots, x^{m})$。因为 A 中任意子集的凸组合都可以看作这 m 个点的凸组合，所以由定理1.1.2可得

$$\mathrm{co}(\boldsymbol{x}^{(1)}, \cdots, \boldsymbol{x}^{(m)})=\left\{\sum_{i=1}^{m} \lambda_i \boldsymbol{x}^{(i)} \middle| \lambda_i \geqslant 0, \quad \sum_{i=1}^{m} \lambda_i = 1\right\}. \tag{1.1-2}$$

由此可引出以下定义.

定义 1.1.4 有限点集$A=\{\boldsymbol{x}^{(1)}, \cdots, \boldsymbol{x}^{(m)}\}$的凸包$\mathrm{co}(\boldsymbol{x}^{(1)}, \cdots, \boldsymbol{x}^{(m)})$称为由$\boldsymbol{x}^{(1)}, \cdots, \boldsymbol{x}^{(m)}$生成的**凸多面体**(convex polyhedron). $\boldsymbol{x}^{(1)}$, …, $\boldsymbol{x}^{(m)}$称为凸多面体的**生成元**.

定义 1.1.5 设$C=\mathrm{co}(\boldsymbol{a}^{(1)}, \cdots, \boldsymbol{a}^{(m)})$, 而$\boldsymbol{a}^{(1)}$, …, $\boldsymbol{a}^{(m)}$中任何一个都不能表示为其余生成元的凸组合，则称$\boldsymbol{a}^{(1)}$, …, $\boldsymbol{a}^{(m)}$为**凸多面体的顶点**(vertex).

在图1-2中点$\boldsymbol{a}^{(1)}, \boldsymbol{a}^{(2)}, \boldsymbol{a}^{(3)}$, $\boldsymbol{a}^{(4)}$是凸多面体C的顶点， 而点$\boldsymbol{a}^{(5)}$不是顶点.

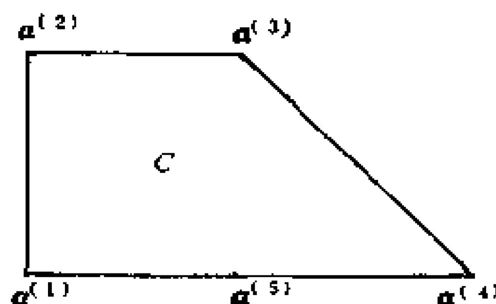


图 1-2

定理 1.1.4 一个凸多面体的顶点集是唯一确定的.

证 设某一凸多面体有两组顶点：$\boldsymbol{a}^{(1)}$, …, $\boldsymbol{a}^{(n)}$和$\boldsymbol{b}^{(1)}$, …, $\boldsymbol{b}^{(m)}$. 因为$\boldsymbol{a}^{(1)}$, …, $\boldsymbol{a}^{(n)}$为一组顶点，所以$\forall j\in\{1, \cdots, m\}$均有

$$\boldsymbol{b}^{(j)} = \sum_{i=1}^{n} \lambda_{ij}\boldsymbol{a}^{(i)}, \quad \lambda_{ij}\geqslant 0, \quad \sum_{i=1}^{n} \lambda_{ij} = 1. \tag{1.1-3}$$

又因为$\boldsymbol{b}^{(1)}$, …, $\boldsymbol{b}^{(m)}$为一组顶点，所以$\forall i\in\{1, \cdots, n\}$均有

$$\boldsymbol{a}^{(i)} = \sum_{t=1}^{m} \mu_{it}\boldsymbol{b}^{(t)}, \quad \mu_{it}\geqslant 0, \quad \sum_{t=1}^{m} \mu_{it} = 1. \tag{1.1-4}$$

将$\boldsymbol{a}^{(i)}$的表达式代入式(1.1-3)，整理之得

$$\begin{aligned}\boldsymbol{b}^{(j)} = &(\lambda_{1j}\mu_{11}+\cdots+\lambda_{nj}\mu_{n1})\boldsymbol{b}^{(1)} \\ &+\cdots\cdots\cdots\cdots \\ &+(\lambda_{1j}\mu_{1m}+\cdots+\lambda_{nj}\mu_{nm})\boldsymbol{b}^{(m)}.\end{aligned}$$

在上式中等号右端$b^{(1)}$, …, $b^{(m)}$的系数非负，且和为1，而且等式右端$b^{(j)}$的系数必为1．否则，$b^{(j)}$就可以表示为这组顶点中其余元素的凸组合．这与$b^{(1)}$, …, $b^{(m)}$为一组顶点矛盾．于是有

$$\lambda_{1j}\mu_{1j}+\cdots+\lambda_{nj}\mu_{nj}=1. \tag{1.1-5}$$

倘若 $\forall k\in\{1, \cdots, n\}$均有$\mu_{kj}<1$，则

$$\lambda_{1j}\mu_{1j}+\cdots+\lambda_{nj}\mu_{nj}<\lambda_{1j}+\cdots+\lambda_{nj}=1,$$

这与式(1.1-5)矛盾．所以必有某个k，使$u_{kj}=1$．由式(1.1-4)得$u_{kt}=0$，$t\neq j$，于是得$a^{(k)}=b^{(j)}$．因而

$$b^{(j)}\in\{a^{(1)}, \cdots, a^{(n)}\}, \ \forall j\in\{1, \cdots, m\}.$$

同理可得

$$a^{(i)}\in\{b^{(1)}, \cdots, b^{(m)}\}, \ \forall i\in\{1, \cdots, n\}.$$

综上所述得

$$\{a^{(1)}, \cdots, a^{(n)}\}=\{b^{(1)}, \cdots, b^{(m)}\}.$$

§1.2 仿射集·仿射包·单纯形

在线性代数中线性子空间与向量组的线性相关性都是重要内容．在凸分析中和它们对应的内容有仿射集与点的仿射相关性．

1.2.1 仿射集与仿射包

定义 1.2.1 设L是V的一个线性子空间，$a\in V$，则L沿a的平移$L+a$称为V的一个**仿射集**（affine set）．

仿射集$L+a$的维数等于线性子空间L的维数，即 $\dim(L+a)=\dim(L)$．

定义 1.2.2 设$A\subset V$．V中所有包含A的仿射集之交称为**A的仿射包**（affine hull），记为aff(A)．

A的维数定义为aff(A)的维数，即

$$\dim(A)=\dim(\mathrm{aff}(A)).$$

定义 1.2.3 设$x^{(1)}$，…，$x^{(k)}\in V$，$\lambda_i\in\mathscr{R}$，$\sum_{i=1}^{k}\lambda_i=1$．则称$x=\sum_{i=1}^{k}\lambda_i x^{(i)}$为$x^{(1)}$，…，$x^{(k)}$的仿射组合．

关于仿射集与仿射包有以下重要性质．

定理 1.2.1 设$M\subset V$，则M是仿射集等价于M包含通过任意两点x，$y\in M$的直线，即

$$\lambda x+(1-\lambda)y\in M,\quad \forall\lambda\in\mathscr{R}.$$

证 1° 设M是仿射集，即$M=L+a$，其中$a\in V$，$L\subset V$是某个线性子空间．

再设x，$y\in M$，即存在$x^{(1)}$，$y^{(1)}\in L$，使$x=x^{(1)}+a$，$y=y^{(1)}+a$，因而$\forall\lambda\in\mathscr{R}$有

$$\lambda x+(1-\lambda)y=\lambda x^{(1)}+(1-\lambda)y^{(1)}+a,$$

而 $z=\lambda x^{(1)}+(1-\lambda)y^{(1)}\in L$，所以$\lambda x+(1-\lambda)y\in M$．

2° 假设 $\forall x$，$y\in M$，$\forall\lambda\in\mathscr{R}$均有

$$\lambda x+(1-\lambda)y\in M.$$

设$a\in M$，令$L=M-a$，再设$\bar{x}$，$\bar{y}\in L$，即存在$x^{(1)}$，$y^{(1)}\in M$，使$\bar{x}=x^{(1)}-a$，$\bar{y}=y^{(1)}-a$，$\forall\lambda\in\mathscr{R}$均有

$$\lambda\bar{x}=\lambda(x^{(1)}-a)=\lambda x^{(1)}+(1-\lambda)a-a.$$

因为 $\lambda x^{(1)}+(1-\lambda)a\in M$，所以$\lambda\bar{x}\in L$，$\forall\lambda\in\mathscr{R}$．

另一方面有

$$\frac{1}{2}\bar{x}+\frac{1}{2}\bar{y}=\frac{1}{2}x^{(1)}+\left(1-\frac{1}{2}\right)y^{(1)}-a,$$

由假设及$x^{(1)}$，$y^{(1)}\in M$可得$\frac{1}{2}x^{(1)}+\left(1-\frac{1}{2}\right)y^{(1)}\in M$，

所以$\frac{1}{2}\bar{x}+\frac{1}{2}\bar{y}\in L$．于是

$$\overline{x}+\overline{y}=2\left(\frac{1}{2}\overline{x}+\frac{1}{2}\overline{y}\right)\in L.$$

综上所述，L是线性子空间．于是，$M=L+a$是仿射集．

以上证明是在$M\neq\Phi$的条件下进行的，若$M=\Phi$，则命题显然正确．

推论1 设M_1，…，$M_k\subset V$均为仿射集，则交集$M=\bigcap_{i=1}^{k}M_i$是仿射集．

证 设$\boldsymbol{x}$，$\boldsymbol{y}\in M$，$\lambda\in\mathscr{R}$．因为M_i为仿射集，而且$\boldsymbol{x},\boldsymbol{y}\in M_i$，$i=1$，…，$k$，所以

$$\lambda\boldsymbol{x}+(1-\lambda)\boldsymbol{y}\in M_i,\quad i=1,\ \cdots,\ k.$$

因而 $\lambda\boldsymbol{x}+(1-\lambda)\boldsymbol{y}\in M$．于是$M$为仿射集．

注 若M为单点集$\{\boldsymbol{x}\}$，则通过M中任意两点$\boldsymbol{x}$，$\boldsymbol{y}$的直线$\lambda\boldsymbol{x}+(1-\lambda)\boldsymbol{y}$ 退化为一点$\boldsymbol{x}$．

推论2 若M是仿射集，则M必为凸集．

定理 1.2.2 设$A\subset V$，则A的仿射包aff(A)等于A中元素的有限仿射组合的全体．

（证明从略，可参阅定理1.1.2的证明）

有限点集$A=\{\boldsymbol{x}^{(1)},\ \cdots,\ \boldsymbol{x}^{(k)}\}$的仿射包记为$\mathrm{aff}(\boldsymbol{x}^{(1)},\ \cdots\boldsymbol{x}^{(k)})$．若$A=\{\boldsymbol{a}\}$，则$\mathrm{aff}(A)=\mathrm{aff}(\boldsymbol{a})=\{\boldsymbol{a}\}$．若$A=\{\boldsymbol{a},\ \boldsymbol{b}\}$，而且$\boldsymbol{a}\neq\boldsymbol{b}$，则$\mathrm{aff}(\boldsymbol{a},\ \boldsymbol{b})$是通过$\boldsymbol{a}$，$\boldsymbol{b}$的直线：

$$\lambda\boldsymbol{a}+(1-\lambda)\boldsymbol{b},\quad\forall\lambda\in\mathscr{R}.$$

若$A=\{\boldsymbol{a},\ \boldsymbol{b},\ \boldsymbol{c}\}$，而且$\boldsymbol{a}$，$\boldsymbol{b}$，$\boldsymbol{c}$不共线，则$\mathrm{aff}(\boldsymbol{a},\ \boldsymbol{b},\ \boldsymbol{c})$是由$\boldsymbol{a}$，$\boldsymbol{b}$，$\boldsymbol{c}$三点决定的平面：

$$\lambda\boldsymbol{a}+\beta\boldsymbol{b}+(1-\lambda-\beta)\boldsymbol{c},\quad\forall\lambda,\ \beta\in\mathscr{R}.$$

一般地有（依定理1.2.2）

$$\mathrm{aff}(\boldsymbol{x}^{(1)},\ \cdots,\ \boldsymbol{x}^{(m)})=\left\{\sum_{i=1}^{m}\lambda_i\boldsymbol{x}^{(i)}\ \middle|\ \lambda_i\in\mathscr{R},\quad\sum_{i=1}^{m}\lambda_i=1\right\}.\tag{1.2-1}$$

1.2.2 仿射无关与仿射相关

定义 1.2.4 设$x^{(1)}, \cdots, x^{(m)} \in V$.若$\dim(\mathrm{aff}(x^{(1)}, \cdots x^{(m)}))=m-1$，则称$x^{(1)}, \cdots, x^{(m)}$**仿射无关**(affinely independent).反之，若$\dim(\mathrm{aff}(x^{(1)}, \cdots, x^{(m)}))<m-1$，则称$x^{(1)}, \cdots x^{(m)}$仿射相关(affinely dependent).

由定义可知，在$\mathscr{R}^n$中，若$x^{(1)}, \cdots, x^{(m)}$仿射无关，则$m\leqslant n+1$.

关于仿射无关，有以下重要定理.

定理 1.2.3 设$x^{(1)}, \cdots, x^{(m)} \in V$，则下列条件等价：

1° $x^{(1)}, \cdots, x^{(m)}$仿射无关；

2° 对任一j，向量组$\{x^{(i)}-x^{(j)} \mid i\neq j\}$线性无关；

3° 若

$$\begin{cases}\alpha_1 x^{(1)}+\cdots+\alpha_m x^{(m)}=0\\ \alpha_1+\cdots+\alpha_m=0,\ \alpha_1, \cdots, \alpha_m\in\mathscr{R}\end{cases} \tag{1.2-2}$$

则 $\alpha_1=\cdots=\alpha_m=0$.

证 要证明三个条件等价，只需证明$1^\circ\Longleftrightarrow 2^\circ$($1^\circ$与$2^\circ$等价)，$2^\circ\Longleftrightarrow 3^\circ$.

先证 $2^\circ\Longleftrightarrow 3^\circ$.

假设2°成立．由式(1.2-2)可得

$$\alpha_1 x^{(1)}+\cdots+\alpha_m x^{(m)}=(\alpha_1+\cdots+\alpha_m)x^{(j)},\ \forall j.$$

由此可得

$$\sum_{i\neq j}\alpha_i(x^{(i)}-x^{(j)})=0$$

因为$\{x^{(i)}-x^{(j)} \mid i\neq j\}$线性无关，所以$\alpha_i=0$，$i\neq j$．再由式(1.2-2)得$\alpha_j=0$．条件$3^\circ$成立.

假设3°成立，任取一$j\in\{1, \cdots, m\}$，令

$$\sum_{i\neq j}\alpha_i(x^{(i)}-x^{(j)})=0,\ \alpha_1, \cdots, \alpha_m\in\mathscr{R}.$$

由此得

$$\sum_{i\neq j}\alpha_i x^{(i)}=\Big(\sum_{i\neq j}\alpha_i\Big)x^{(j)}.$$

再令 $\alpha_j=-\sum_{i\neq j}\alpha_i$，则得式(1.2-2)。由假设可得$\alpha_1=\cdots=\alpha_m=0$。因而$\{x^{(i)}-x^{(j)}\mid i\neq j\}$线性无关．条件2°成立．

再证1°$\Longleftrightarrow$2°．

假设2°成立，因为$x^{(2)}-x^{(1)}$，…，$x^{(m)}-x^{(1)}$线性无关，所以由它们生成的子空间L的维数为$m-1$．而$\text{aff}(x^{(1)},\cdots,x^{(m)})=L+x^{(1)}$，所以$\dim(\text{aff}(x^{(1)},\cdots,x^{(m)}))=m-1$．命题1°成立．

反之，若对某一j，不妨设$j=1$，向量组$x^{(2)}-x^{(1)}$，…，$x^{(m)}-x^{(1)}$线性相关，则由它们生成的子空间L的维数小于$m-1$．而$\text{aff}(x^{(1)},\cdots,x^{(m)})=L+x^{(1)}$，$\dim(\text{aff}(x^{(1)},\cdots,x^{(m)}))=\dim L<m-1$，于是$x^{(1)}$，…，$x^{(m)}$仿射相关．所以，若1°成立，则2°必成立．

推论1 若$x^{(1)}$，…，$x^{(m)}$仿射无关，则$\forall x\in\text{aff}(x^{(1)},\cdots,x^{(m)})$都可唯一地表示为$x^{(1)}$，…$x^{(m)}$的仿射组合．

推论2 $x^{(1)}$，…，$x^{(m)}$仿射相关等价于存在不全为零的数α_1，…，$\alpha_m\in\mathscr{R}$使式(1.2-2)成立．

推论3 若$x^{(1)}$，…，$x^{(m)}$仿射相关，则其中必有一个可以表示为其余$m-1$个点的仿射组合．

证 因为$x^{(1)}$，…，$x^{(m)}$仿射相关，所以由定理1.2.3可知，对于某个j，不妨设$j=1$，向量组$x^{(2)}-x^{(1)}$，…，$x^{(m)}-x^{(1)}$线性相关．因而存在不全为零的数$\alpha_i\in\mathscr{R}$，使$\sum_{i=2}^{m}\alpha_i(x^{(i)}-x^{(1)})=0$．不妨设$\alpha_m\neq0$．以$\alpha_m$除之，并将$\alpha_i/\alpha_m$记为$\lambda_i$．于是可得

$$x^{(m)}=x^{(1)}-\sum_{i=2}^{m-1}\lambda_i(x^{(i)}-x^{(1)})$$

$$=\left(1+\sum_{i=2}^{m-1}\lambda_i\right)x^{(1)}-\sum_{i=2}^{m-1}\lambda_i x^{(i)}$$

其中$x^{(1)}$, …, $x^{(m-1)}$的系数之和为1，即$x^{(m)}$表示为$x^{(1)}$, …, $x^{(m-1)}$的仿射组合.

推论4 1° 在$\mathscr{R}^n$中存在$n+1$个仿射无关点；

2° 设$x^{(1)}$, …, $x^{(k)}\in\mathscr{R}^n$仿射无关，而且$k<n+1$，则可以扩张为包含$n+1$个仿射无关点的仿射集.

证 1° 因为$e^{(1)}-0$,…, $e^{(n)}-0$线性无关，所以0, $e^{(1)}$, …, $e^{(n)}$为仿射无关集，其中$e^{(j)}$为第j个分量为1，其余分量均为零的单位列向量.

2° 因为$x^{(1)}$, …, $x^{(k)}$仿射无关，所以$x^{(2)}-x^{(1)}$, …, $x^{(k)}-x^{(1)}$线性无关，因而存在$a^{(k+1)}$, …, $a^{(n+1)}\in\mathscr{R}^n$，使$x^{(2)}-x^{(1)}$, …, $x^{(k)}-x^{(1)}$, $a^{(k+1)}$, …, $a^{(n+1)}$线性无关. 令$x^{(i)}=a^{(i)}+x^{(1)}$, $i=k+1$, …, $n+1$. 再设

$$\begin{cases}\alpha_1 x^{(1)}+\alpha_2 x^{(2)}+\cdots+\alpha_{n+1}x^{(n+1)}=0\\ \alpha_1+\alpha_2+\cdots+\alpha_{n+1}=0\end{cases}$$

由此可得

$$\alpha_1 x^{(1)}+\alpha_2 x^{(2)}+\cdots+\alpha_{n+1}x^{(n+1)}=(\alpha_1+\alpha_2+\cdots+\alpha_{n+1})x^{(1)}$$

因而

$$\alpha_2(x^{(2)}-x^{(1)})+\cdots+\alpha_{n+1}(x^{(n+1)}-x^{(1)})=0.$$

由于$x^{(2)}-x^{(1)}$, …, $x^{(n+1)}-x^{(1)}$线性无关，故$\alpha_2=\cdots=\alpha_{n+1}=0$. 于是$\alpha_1=0$. 由定理1.2.3知$x^{(1)}$, …, $x^{(n+1)}$仿射无关.

1.2.3 单纯形

定义 1.2.5 设$x^{(0)}$, $x^{(1)}$, …, $x^{(k)}\in V$. 若它们仿射无关，则它们的凸包co($x^{(0)}$, $x^{(1)}$, …, $x^{(k)}$)称为一个**k维单纯形**(k-simplex).

零维单纯形是一个点co(a)=\{a\}. 1维单纯形是一条线段

$\text{co}(x^{(0)}, x^{(1)}) = [x^{(0)}, x^{(1)}]$. 2 维单纯形是一个三角形：

$$\left\{\lambda_0 x^{(0)} + \lambda_1 x^{(1)} + \lambda_2 x^{(2)} \,\middle|\, \lambda_i \geqslant 0, \quad \sum_{i=0}^{2} \lambda_i = 1\right\},$$

3 维单纯形是一个四面体：

$$\left\{\sum_{i=0}^{3} \lambda_i x^{(i)} \,\middle|\, \lambda_i \geqslant 0, \quad \sum_{i=0}^{3} \lambda_i = 1\right\}.$$

定理 1.2.4 若$C = \text{co}(x^{(0)}, x^{(1)}, \cdots, x^{(k)})$是$V$中一个$k$维单纯形，则

1° $x^{(0)}, x^{(1)}, \cdots, x^{(k)}$是$C$的顶点；

2° 每个$x \in C$都能唯一地表示为$x^{(0)}, x^{(1)}, \cdots, x^{(k)}$的一个凸组合.

证 1° 倘若不然，则有$x^{(i)}$，不妨设 $i = 0$，可以表示为其余各点的凸组合：

$$x^{(0)} = \sum_{i=1}^{k} \lambda_i x^{(i)}, \quad \lambda_i \geqslant 0, \quad \sum_{i=1}^{k} \lambda_i = 1.$$

由此可得

$$x^{(0)} - \sum_{i=1}^{k} \lambda_i x^{(i)} = 0, \quad 1 - \lambda_1 - \cdots \lambda_k = 0.$$

由定理1.2.3的推论 2 可知，$x^{(0)}, x^{(1)}, \cdots, x^{(k)}$仿射相关.
这与C为一个k维单纯形矛盾. 所以$x^{(0)}, x^{(1)}, \cdots, x^{(k)}$是 C 的顶点.

2° 根据式(1.1-2)，只需证明唯一性. 设对于$x \in C$有

$$x = \sum_{i=0}^{k} \lambda_i x^{(i)}, \quad \lambda_i \geqslant 0, \quad \sum_{i=0}^{k} \lambda_i = 1,$$

$$x = \sum_{i=0}^{k} \mu_i x^{(i)}, \quad \mu_i \geqslant 0, \quad \sum_{i=0}^{k} \mu_i = 1.$$

于是有

$$\sum_{i=0}^{k}(\lambda_i-\mu_i)x^{(i)}=0,$$

而且

$$\sum_{i=0}^{k}(\lambda_i-\mu_i)=\sum_{i=0}^{k}\lambda_i-\sum_{i=0}^{k}\mu_i=0.$$

因为$x^{(0)}$，$x^{(1)}$，…，$x^{(k)}$仿射无关，所以根据定理1.2.3得$\lambda_i-\mu_i=0$，$i=0, 1, \cdots, k$，因而 $\lambda_i=\mu_i$，$i=0, 1, \cdots, k$。唯一性得证。

§1.3 代数内部与代数闭包·凸代数体

在本节中首先引入集合A的代数内部和代数闭包的概念，再研究它们的性质。然后引入凸代数体的概念与Minkowski 泛函，并以后者为工具研究凸代数体。

1.3.1 代数内部与代数闭包

定义 1.3.1 设$A\subset V$. A的**代数内部**（algebraic interior）A^i由所有这样的点x组成：对于V中每一条通过x的直线m，它与A的交$m\cap A$包含着以x为内点的线段。

定义 1.3.2 设$A\subset V$. A的**代数闭包**（algebraic closure）A^a由所有这样的点$x\in V$组成：对于它存在$c\in A$，使$[c, x)\subset A$（见图1-3）。

显然$A^i\subset A\subset A^a$。

对于A的代数内部A^i与代数闭包A^a有以下性质（见图1-3）：

性质1 $x\in A^i \Longleftrightarrow \forall v\in V$，存在$\delta>0$使

$$(x-\delta v, x+\delta v)\subset A$$

$$\Longleftrightarrow \forall v\in V,\ 存在\delta>0使[x, x+\delta v)\subset A.$$

性质2 若$A\subset V$，$b\in V$，则$(A+b)^i=A^i+b$，

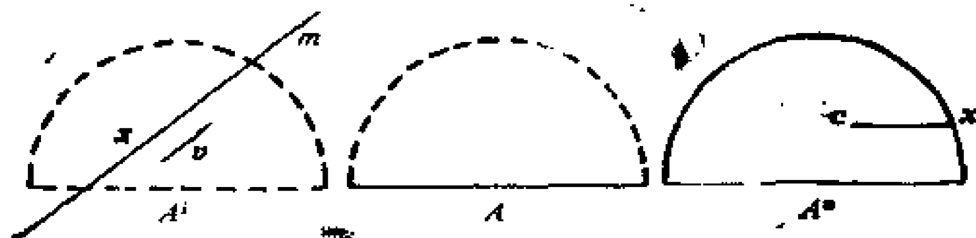


图 1-3

$$(A+\boldsymbol{b})^a=A^a+\boldsymbol{b}.$$

定理 1.3.1 设$A\subset V$为凸集，则A^i与A^a均为凸集.

证 1° 设$\boldsymbol{x}$，$\boldsymbol{y}\in A^i$，$\lambda\in(0, 1)$，则$\boldsymbol{z}=\lambda\boldsymbol{x}+(1-\lambda)\boldsymbol{y}\in A$. 依性质 1，$\forall\boldsymbol{v}\in V$，存在$\delta>0$使

$$[\boldsymbol{x}, \boldsymbol{x}+\delta\boldsymbol{v}]\subset A, [\boldsymbol{y}, \boldsymbol{y}+\delta\boldsymbol{v}]\subset A.$$

因而$\boldsymbol{x}+\delta\boldsymbol{v}$，$\boldsymbol{y}+\delta\boldsymbol{v}\in A$. 因为$A$为凸集，所以

$$\boldsymbol{z}+\delta\boldsymbol{v}=\lambda(\boldsymbol{x}+\delta\boldsymbol{v})+(1-\lambda)(\boldsymbol{y}+\delta\boldsymbol{v})\in A,$$

于是$[\boldsymbol{z}, \boldsymbol{z}+\delta\boldsymbol{v}]\subset A$. 再由性质 1 得$\boldsymbol{z}\in A^i$，所以$A^i$是凸集.

2° 设$\boldsymbol{x}$，$\boldsymbol{y}\in A^a$，$\lambda\in(0, 1)$，$\boldsymbol{z}=\lambda\boldsymbol{x}+(1-\lambda)\boldsymbol{y}$. 根据定义1.3.2存在$\boldsymbol{c}_1$，$\boldsymbol{c}_2\in A$，使$[\boldsymbol{c}_1, \boldsymbol{x})\subset A$，$[\boldsymbol{c}_2, \boldsymbol{y})\subset A$. 令$\boldsymbol{c}=\lambda\boldsymbol{c}_1+(1-\lambda)\boldsymbol{c}_2$，则$\boldsymbol{c}\in A$. 再设$\mu\in(0, 1)$，则

$$\mu\boldsymbol{c}_1+(1-\mu)\boldsymbol{x}\in A, \mu\boldsymbol{c}_2+(1-\mu)\boldsymbol{y}\in A,$$

又因为

$$\begin{aligned}&\mu\boldsymbol{c}+(1-\mu)\boldsymbol{z}\\&=\mu[\lambda\boldsymbol{c}_1+(1-\lambda)\boldsymbol{c}_2]+(1-\mu)[\lambda\boldsymbol{x}+(1-\lambda)\boldsymbol{y}]\\&=\lambda[\mu\boldsymbol{c}_1+(1-\mu)\boldsymbol{x}]+(1-\lambda)[\mu\boldsymbol{c}_2+(1-\mu)\boldsymbol{y}]\end{aligned}$$

所以 $\mu\boldsymbol{c}+(1-\mu)\boldsymbol{z}\in A$，$\forall\mu\in(0, 1)$，因而$[\boldsymbol{c}, \boldsymbol{z})\subset A$. 依定义1.3.2 $\boldsymbol{z}\in A^a$. 故A^a为凸集.

定理 1.3.2 设$A\subset V$是凸集，$\boldsymbol{x}\in A^i$，$\boldsymbol{y}\in A^a$，则$[\boldsymbol{x}, \boldsymbol{y})\subset A^i$.

证 设$\boldsymbol{z}=\lambda\boldsymbol{x}+(1-\lambda)\boldsymbol{y}$，其中$\lambda\in(0, 1)$.

1° 设$\boldsymbol{y}\in A$，因而$\boldsymbol{z}\in A$. 因为$\boldsymbol{x}\in A^i$，所以$\forall\boldsymbol{v}\in V$都存在

$\delta>0$使$[x, x+\delta v)\subset A$（根据性质1），因而$x+\delta v\in A$. 再由A的凸性得

$$z+\lambda\delta v=\lambda(x+\delta v)+(1-\lambda)y\in A.$$

于是$[z, z+\lambda\delta v]\subset A$, 依性质1 $z\in A^i$, 因而$[x, y)\subset A^i$.

2° 设$y\in A^a\setminus A$. 因为$y\in A^a$,所以存在$u\in A$, 使$[u, y)\subset A$.

假设x, y, u三点共线，不妨设$u\in[x, y)$. 任取$y^{(1)}\in(u, y)\subset A$（见图1-4, a），由1°可得$[x, y^{(1)})\subset A^i$, 由$y^{(1)}$的任意性得$[x, y)\subset A^i$.

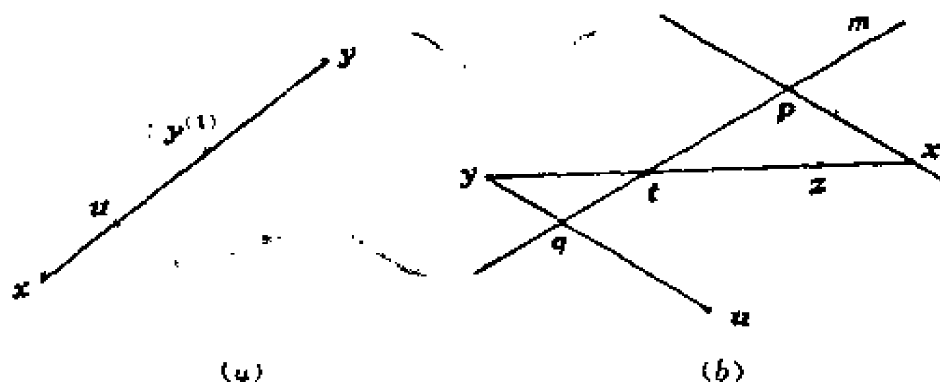


图 1-4

假设x, y, u三点不共线，则取$t\in(y, z)$. 因而$z\in(x, t)$. 由$x\in A^i$可知，存在充分小的正数$\delta>0$使$p=x+\delta(y-u)\in A$. 令m为通过点p与t的直线. 因为$\delta>0$充分小,所以直线m与$[u, y)$交于一点$q\in A$（见图1-4, b）. 再由A的凸性得$t\in A$. 于是由1°得$[x, t)\subset A^i$, 因而$z\in A^i$. 再由z的任意性得$[x, y)\subset A^i$.

推论 设$A\subset V$为凸集，则$(A^i)^i=A^i$.

证 由定义1.3.1得$(A^i)^i\subset A^i$.

设$x\in A^i$. 由性质1可知，$\forall v\in V$均存在$\delta>0$，使$[x, x+\delta v)\subset A$. 因而$x+\delta v\in A\subset A^a$. 由定理1.3.2得$[x, x+\delta v)\subset A^i$. 令$\delta_1=\frac{1}{2}\delta$, 则$[x, x+\delta_1 v]\subset A^i$, 因而$x\in(A^i)^i$. 于是得$A^i\subset$

$(A^i)^i$.

综上所述，命题得证.

注 若A不是凸集，则$(A^i)^i$与A^i可能不等. 下面举一反例.

例 设$V=\mathscr{R}^2$，

$$A=\{(x,\ \lambda)|\ \lambda\geqslant x^2\}\cup\{(x,\ \lambda)|\ \lambda=0\}$$
$$\cup\{(x,\ \lambda)|\ \lambda\leqslant -x^2\}\ (见图1\text{-}5),$$

试写出A^i与$(A^i)^i$.

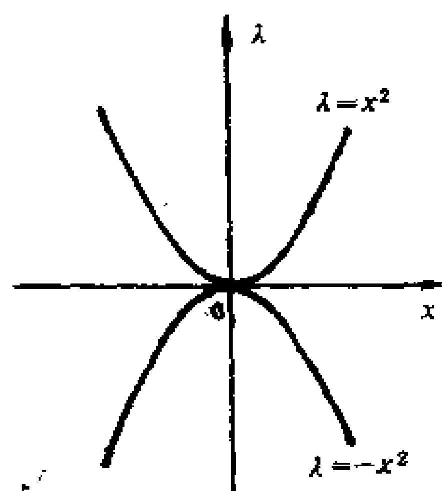


图 1-5

解 $A^i=\{(x,\ \lambda)\ |\lambda>x^2\}$
$\cup\{(0,\ 0)\}$
$\cup\{(x,\ \lambda)|\ \lambda<-x^2\}$,

$(A^i)^i=A^i\setminus\{(0,\ 0)\}\neq A^i$.

定理 1.3.3 设$C,D\subset V$是非空凸集，$C\cap D=\Phi$，$C\cup D=V$，$H=C^a\cap D^a$，则

1° $C^i\cap D^a=\Phi$；$C^a\cap D^i=\Phi$；

2° $C^i\cup D^a=V$；$C^a\cup D^i=V$；

3° H是一个仿射集；

4° 若$\boldsymbol{x}\in C^i$，$\boldsymbol{y}\in D$，则存在唯一点$\boldsymbol{z}\in(\boldsymbol{x},\boldsymbol{y})\cap H$.

证 设$\boldsymbol{x}\in V$.

1° 若$\boldsymbol{x}\in D^a\setminus D$,则存在$\boldsymbol{u}\in D$,使$(\boldsymbol{u},\boldsymbol{x})\subset D$,因而$(\boldsymbol{u},\boldsymbol{x})\cap C=\Phi$，所以$\boldsymbol{x}\overline{\in}C^i$；若$\boldsymbol{x}\in D$，则$\forall\ \boldsymbol{u}\in D$都有$(\boldsymbol{u},\boldsymbol{x})\subset D$，因而$\boldsymbol{x}\overline{\in}C^i$. 所以$C^i\cap C^a=\Phi$. 同理可证另一部分.

2° 若$\boldsymbol{x}\overline{\in}D^a$,则对于任一点$\boldsymbol{u}\in V$都存在一点$\boldsymbol{z}\in(\boldsymbol{u},\boldsymbol{x})\cap C$. 令$\boldsymbol{u}^{(1)}=\boldsymbol{x}+(\boldsymbol{x}-\boldsymbol{u})=2\boldsymbol{x}-\boldsymbol{u}$，必存在$\boldsymbol{z}^{(1)}\in(\boldsymbol{u}^{(1)},\boldsymbol{x})\cap C$. 于是，$\boldsymbol{x}\in\mathrm{int}(\boldsymbol{z},\boldsymbol{z}^{(1)})\subset C$. 再由定义1.3.1得$\boldsymbol{x}\in C^i$. 因而$C^i\cup D^a=V$. 同理可证另一部分.

3° 由式理1.3.1知C^a，D^a均为凸集,因而它们的交H是凸集. 设$\boldsymbol{x},\boldsymbol{y}\in H$，则$(\boldsymbol{x},\boldsymbol{y})\subset H$.若通过$\boldsymbol{x},\boldsymbol{y}$的直线$m$上有一点$\boldsymbol{z}\overline{\in}H$，

则$z \overline{\in} [x,y]$. 因而有$y\in(x,z)$或$x\in(z,y)$. 不妨设$x\in(z,y)$. 由于$z \overline{\in} H$，所以$z\in C^i$或D^i. 不妨设$z\in C^i$. 依定理1.3.2得$(z,y)\subset C^i$. 因而$x\in C^i$. 再由1°知$x \overline{\in} D^a$. 这与$x\in H$矛盾. 所以直线m必含于H中. 依定理1.2.1 H是仿射集.

4° （证明留给读者）

例 1 若$C=V$，则$C^i=C^a=V$.

例 2 若$V=\mathscr{R}^2$，$C=(a,b]$，则$C^i=\Phi$，$C^a=[a,b]$.

例 3 设V为由$\mathscr{R}$上的所有单元多项式所构成的线性空间，C为其中首项系数为正的单元多项式所构成的集合. 试证明$C^a=V$，$C^i=\Phi$.

证 设$y\in V$，即 $y=\sum_{i=0}^{m}\alpha_i x^i$，$\alpha_i$，$x\in\mathscr{R}$，

$i=0, 1, \cdots, m$.

1° 令 $u=\sum_{i=0}^{N}\beta_i x^i$，其中$\beta_i, x\in\mathscr{R}$，$j=0, 1, \cdots, N$，$N>m+1$，$\beta_N>0$，则$u\in C$. 再设 $\alpha_i=0$，$i=m+1, \cdots, N-1$.

当$z\in[u, y)$时，存在$\lambda\in(0, 1]$使

$$
\begin{aligned}
z&=\lambda u+(1-\lambda)y\\
&=\lambda\beta_N x^N+\sum_{i=0}^{N-1}[\lambda\beta_i+(1-\lambda)\alpha_i]x^i\in C.
\end{aligned}
$$

由z的任意性得$[u,y)\subset C$，因而$y\in C^a$. 这意味着$V\subset C^a$，而$C^a\subset V$，所以$C^a=V$.

2° 令 $u=\sum_{i=0}^{N}\mu_i x^i$，$\mu_N=-1$，$\mu_i, x\in\mathscr{R}$，$i=0, 1, \cdots N-1$，$N>m+1$. 再令$\alpha_i=0$，$i=m+1, \cdots, N-1$.

当$z\in[u, y)$时，存在$\lambda\in(0, 1]$使

$$
\begin{aligned}
z &= \lambda u + (1-\lambda) y \\
&= -\lambda x^N + \sum_{i=0}^{N-1} [\lambda\mu_i + (1-\lambda)\alpha_i] x^i \overline{\in} C.
\end{aligned}
$$

因为z是任意的，所以依定义1.3.1$y \overline{\in} C^i$. 再由y的任意性知$C^i = \Phi$.

因代数内部不空的凸集具有更重要的意义，故着重讨论这类凸集：

1.3.2 凸代数体

定义 1.3.3 若$C \subset V$是凸集，且$C^i \neq \Phi$，则称C为**凸代数体**（convex algebraic body）.

设$C \subset V$是一个凸代数体，且$0 \in C^i$. 由性质1知，$\forall x \in V$，存在$\mu > 0$，使$[0, \frac{1}{\mu}x] \subset C$. 所以$\frac{1}{\mu}x \in C$，因而$x \in \mu C$. 于是引入以下定义.

定义 1.3.4 设$C \subset V$是是一个凸代数体，且$0 \in C^i$，则称

$$
p(x) = \inf\{\mu \mid \mu > 0, \ x \in \mu C\} \qquad (1.3\text{-}1)
$$

为关于C的Minkowski泛函（或称为规范函数）.

由定义可知，当$x \in C$时$p(x) \leqslant 1$；当$x \overline{\in} C$时$p(x) \geqslant 1$，因而，若$p(x) < 1$，则$x \in C$.

若V为赋范线性空间，$C \subset V$为单位球，则$p(x)$恰好是点x的范数$\|x\|$（即点x与点0的距离）.

Minkowski泛函$p(x)$具有以下性质：

性质3 $p(x) \geqslant 0$（显然）.

性质4 $p(\lambda x) = \lambda p(x)$，$\forall \lambda \geqslant 0$（正齐次性）.

证 当$\lambda = 0$时上式显然成立. 当$\lambda > 0$时

$$
p(\lambda x) = \inf\{\mu \mid \mu > 0, \ \lambda x \in \mu C\}
$$

$$
= \inf\left\{\lambda \cdot \frac{\mu}{\lambda} \,\middle|\, \mu < 0, \ x \in \frac{\mu}{\lambda} C\right\}
$$

$$=\lambda\inf\{\alpha\,|\,\alpha>0,\ \boldsymbol{x}\in\alpha C\}$$
$$=\lambda p(\boldsymbol{x}),$$

所以命题成立.

性质5 $p(\boldsymbol{x}+\boldsymbol{y})\leqslant p(\boldsymbol{x})+p(\boldsymbol{y}),\ \forall\,\boldsymbol{x},\boldsymbol{y}\in V$（次 加性）.

证 设$p(\boldsymbol{x})=a$，$p(\boldsymbol{y})=b$. 根据$p(\boldsymbol{x})$的定义，$\forall\,\varepsilon>0$ 必有

$$\boldsymbol{x}\in(a+\varepsilon)C,\quad \boldsymbol{y}\in(b+\varepsilon)C,$$

由此得

$$\boldsymbol{x}+\boldsymbol{y}\in(a+\varepsilon)C+(b+\varepsilon)C$$
$$=(a+b+2\varepsilon)C\quad（依\,§\,1.1性质\,3\,）.$$

从而 $p(\boldsymbol{x}+\boldsymbol{y})\leqslant a+b+2\varepsilon=p(\boldsymbol{x})+p(\boldsymbol{y})+2\varepsilon.$

令 $\varepsilon\to0$ 得

$$p(\boldsymbol{x}+\boldsymbol{y})\leqslant p(\boldsymbol{x})+p(\boldsymbol{y}).$$

易知对于赋范线性空间V中的单位球C有以下性质：

$$C^i=\{\boldsymbol{x}\,|\,p(\boldsymbol{x})<1\},\quad C^a=\{\boldsymbol{x}\,|\,p(\boldsymbol{x})\leqslant1\}.$$

定理 1.3.4 设$C\subset V$是一个凸代数体，且$0\in C^i$，$p(\boldsymbol{x})$为关于C的Minkowski泛函. 则

$$C^i=(C^a)^i=\{\boldsymbol{x}\,|\,p(\boldsymbol{x})<1\},\tag{1.3-2}$$
$$C^a=(C^i)^a=\{\boldsymbol{x}\,|\,p(\boldsymbol{x})\leqslant1\}.\tag{1.3-3}$$

证 1° 设$\boldsymbol{x}\in C^a$. 由定义1.3.2知存在$\boldsymbol{z}\in C$使$(\boldsymbol{z},\boldsymbol{x})\subset C$. 即$\forall\,\lambda\in(0,1)$ $\lambda\boldsymbol{x}+(1-\lambda)\boldsymbol{z}=\boldsymbol{z}+\lambda(\boldsymbol{x}-\boldsymbol{z})\in C$. 因此 $p(\boldsymbol{z}+\lambda(\boldsymbol{x}-\boldsymbol{z}))\leqslant1$. 而$\boldsymbol{x}=\boldsymbol{z}+\lambda(\boldsymbol{x}-\boldsymbol{z})+(1-\lambda)(\boldsymbol{x}-\boldsymbol{z})$，所 以

$$p(\boldsymbol{x})\leqslant p(\boldsymbol{z}+\lambda(\boldsymbol{x}-\boldsymbol{z}))+(1-\lambda)p(\boldsymbol{x}-\boldsymbol{z})$$
$$\leqslant1+(1-\lambda)p(\boldsymbol{x}-\boldsymbol{z}).$$

令$\lambda\to1$得$p(\boldsymbol{x})\leqslant1$，因而$C^a\subset\{\boldsymbol{x}\,|\,p(\boldsymbol{x})\leqslant1\}$.

2° 设$\boldsymbol{x}\in(C^a)^i$. 依定义1.3.1存在$\lambda>1$使$\lambda\boldsymbol{x}\in C^a$. 由1°得 $\lambda p(\boldsymbol{x})=p(\lambda\boldsymbol{x})\leqslant1$，所以 $p(\boldsymbol{x})<1$，因而$(C^a)^i\subset\{\boldsymbol{x}\,|\,p(\boldsymbol{x})<1\}$.

3° 设$\boldsymbol{x},\boldsymbol{y}\in V$，$p(\boldsymbol{x})<1$，因而$\boldsymbol{x}\in C$.对于充分小的$\delta>0$有

$$p(\boldsymbol{x}+\delta\boldsymbol{y})\leqslant p(\boldsymbol{x})+\delta p(\boldsymbol{y})<1,$$

所以$\boldsymbol{x}+\delta\boldsymbol{y}\in C$，因而$[\boldsymbol{x},\boldsymbol{x}+\delta\boldsymbol{y}]\subset C$. 由$\boldsymbol{y}$的任意性与性质1得$\boldsymbol{x}\in C^i$，因而$\{\boldsymbol{x}\mid p(\boldsymbol{x})<1\}\subset C^i$.

因为$C^i\subset(C^a)^i$，所以由2°与3°可得式(1.3-2).

4° 设$\boldsymbol{x}\in V$，$p(\boldsymbol{x})=1$. 当$\boldsymbol{z}\in[\boldsymbol{0},\boldsymbol{x})$时，$\boldsymbol{z}=\lambda\boldsymbol{x}$，$0\leqslant\lambda<1$. $p(\boldsymbol{z})=p(\lambda\boldsymbol{x})=\lambda p(\boldsymbol{x})<1$. 因而$[\boldsymbol{0},\boldsymbol{x})\subset C$. 依定义1.3.2 $\boldsymbol{x}\in C^a$. 又因为$\boldsymbol{0}\in C^i$，所以由定理1.3.2得$[\boldsymbol{0},\boldsymbol{x})\subset C^i$，因而$\boldsymbol{x}\in(C^i)^a$. 结合3°得$\{\boldsymbol{x}\mid p(\boldsymbol{x})\leqslant 1\}\subset(C^i)^a$.

因为$(C^i)^a\subset C^a$，所以结合1°与4°得式(1.3-3).

推论 在定理1.3.4的条件下有$(C^a)^a=C^a$.

证 1° 依定义知$C^a\subset(C^a)^a$.

2° 设$\boldsymbol{x}\in(C^a)^a$. 若$\boldsymbol{x}\overline{\in}C^a$，则存在$\boldsymbol{z}\in C^a$，使$[\boldsymbol{z},\boldsymbol{x})\subset C^a$，即$\forall\lambda\in[0,1)$均有

$$\lambda\boldsymbol{x}+(1-\lambda)\boldsymbol{z}=\boldsymbol{z}+\lambda(\boldsymbol{x}-\boldsymbol{z})\in C^a,$$

因而$p(\boldsymbol{z}+\lambda(\boldsymbol{x}-\boldsymbol{z}))\leqslant 1$. 又因为

$$\boldsymbol{x}=\boldsymbol{z}+\lambda(\boldsymbol{x}-\boldsymbol{z})+(1-\lambda)(\boldsymbol{x}-\boldsymbol{z})$$

所以有

$$\begin{aligned}p(\boldsymbol{x})&\leqslant p(\boldsymbol{z}+\lambda(\boldsymbol{x}-\boldsymbol{z}))+(1-\lambda)p(\boldsymbol{x}-\boldsymbol{z})\\&\leqslant 1+(1-\lambda)p(\boldsymbol{x}-\boldsymbol{z}).\end{aligned}$$

令$\lambda\to 1$得$p(\boldsymbol{x})\leqslant 1$. 这与定理1.3.4矛盾. 所以$\boldsymbol{x}\in C^a$. 因而$(C^a)^a\subset C^a$.

综合1°与2°得$(C^a)^a=C^a$.

*§1.4 线性拓扑空间的凸集

从本节起用E表示$\mathscr{R}$上的线性拓扑空间. 为方便起见，假设E中包含的点多于一个.

本节首先研究凸集的内部与闭包的性质，然后再研究凸集的内部与代数内部、凸集的闭包与代数闭包之间的关系.

定理1.4.1 设$C\subset E$是凸集.

1° 若$\boldsymbol{x}\in \mathrm{int}(C),\boldsymbol{y}\in \overline{C}$，则$(\boldsymbol{x},\boldsymbol{y})\subset \mathrm{int}(C)$;

2° C的内部$\mathrm{int}(C)$是凸集;

3° C的闭包$\overline{C}$是凸集.

证 1° 设$\boldsymbol{x}\in \mathrm{int}(C),\boldsymbol{y}\in \overline{C}$，$\lambda\in(0,1)$，$\boldsymbol{z}=\lambda\boldsymbol{x}+(1-\lambda)\boldsymbol{y}$. 再设$U\subset C$是$\boldsymbol{x}$的一个邻域. 令

$$W=\frac{1}{1-\lambda}(\boldsymbol{z}-\lambda U),$$

则W是$\boldsymbol{y}$的一个邻域. 因为$\boldsymbol{y}\in \overline{C}$，所以存在$\boldsymbol{c}\in W\cap C$. 可设

$$\boldsymbol{c}=\frac{1}{1-\lambda}(\boldsymbol{z}-\lambda\boldsymbol{u}),$$

其中$\boldsymbol{u}\in U$. 由此得$\boldsymbol{z}=\lambda\boldsymbol{u}+(1-\lambda)\boldsymbol{c}$. 因为$U$是$\boldsymbol{u}$的邻域，所以

$$Z=\lambda U+(1-\lambda)\boldsymbol{c}$$

是$\boldsymbol{z}$的邻域. 又因为$\boldsymbol{c}\in C$，$U\subset C$，所以$Z\subset C$. 因而得$\boldsymbol{z}\in \mathrm{int}(C)$. 再由$\boldsymbol{z}$的任意性得$(\boldsymbol{x},\boldsymbol{y})\subset \mathrm{int}(C)$.

2° 设$\boldsymbol{x},\boldsymbol{y}\in \mathrm{int}(C)$. 由1°得$(\boldsymbol{x},\boldsymbol{y})\subset \mathrm{int}(C)$，因而$\mathrm{int}(C)$是凸集.

3° 设$\boldsymbol{x}^{(0)}$, $\boldsymbol{y}^{(0)}\in\overline{C}$，$\lambda\in(0,1)$，$\boldsymbol{z}^{(0)}=\lambda\boldsymbol{x}^{(0)}+(1-\lambda)\boldsymbol{y}^{(0)}$. 再设$U$是$\boldsymbol{z}^{(0)}$的任一邻域. 由映射$(\boldsymbol{x},\boldsymbol{y})\to\lambda\boldsymbol{x}+(1-\lambda)\boldsymbol{y}$的连续性知，存在$\boldsymbol{x}^{(0)}$的邻域$U_1$与$\boldsymbol{y}^{(0)}$的邻域$U_2$使$\lambda U_1+(1-\lambda)U_2\subset U$. 再由$\boldsymbol{x}^{(0)}$，$\boldsymbol{y}^{(0)}\in\overline{C}$知，存在$\boldsymbol{c}_1\in U_1\cap C$，$\boldsymbol{c}_2\in U_2\cap C$. 从而

$$\lambda\boldsymbol{c}_1+(1-\lambda)\boldsymbol{c}_2\in(\lambda U_1+(1-\lambda)U_2)\cap C\subset U\cap C$$

于是$\boldsymbol{z}^{(0)}\in\overline{C}$，所以$\overline{C}$是凸集.

定理 1.4.2 若$A\subset E$是开集，则$\mathrm{co}(A)$也是开集

证 因为A是开集，且$A\subset \mathrm{co}(A)$，所以

$$A=\text{int}(A)\subset\text{int}(\text{co}(A)).$$

由定理1.4.1知int(co(A))是凸集，所以

$$\text{co}(A)\subset\text{int}(\text{co}(A)).$$

又因为int(co(A))$\subset$co(A)，所以

$$\text{co}(A)=\text{int}(\text{co}(A)),$$

因而co(A)是开集.

定义 1.4.1 设$A\subset E$.E中所有包含A的闭凸集的交称为A**的闭凸包**，记为$\overline{\text{co}}(A)$.

显然，闭凸包$\overline{\text{co}}(A)$是E中包含A的最小闭凸集.

定理 1.4.3 设$A\subset E$，则$\overline{\text{co}}(A)=\overline{\text{co}(A)}$.

证 因为$A\subset\text{co}(A)\subset\overline{\text{co}(A)}$，而且$\overline{\text{co}(A)}$是闭凸集，所以$\overline{\text{co}}(A)\subset\overline{\text{co}(A)}$.

反之，由定义可得$\text{co}(A)\subset\overline{\text{co}}(A)$，所以$\overline{\text{co}(A)}\subset\overline{\text{co}}(A)$. 于是命题得证.

下面研究凸集的内部与代数内部、凸集的闭包与代数闭包之间的关系.

定理 1.4.4 设$A\subset E$. 则

$$\text{int}(A)\subset A^i,\ A^a\subset\overline{A}.$$

证 1° 设$\boldsymbol{x}\in\text{int}(A)$，$\boldsymbol{v}\in E$. 令$U$为$\boldsymbol{x}$的一个邻域，且$U\subset A$. 存在$\delta>0$使$\forall\lambda\in[0,\delta]$均有$\boldsymbol{x}+\lambda\boldsymbol{v}\in U$，即$[\boldsymbol{x},\boldsymbol{x}+\delta\boldsymbol{v}]\subset U\subset A$. 由§1.3性质1得$\boldsymbol{x}\in A^i$，因而$\text{int}(A)\subset A^i$.

2° 设$\boldsymbol{x}\in A^a$. 若$\boldsymbol{x}\in A$，则$\boldsymbol{x}\in\overline{A}$.若$\boldsymbol{x}\notin A$，则依定义1.3.2存在$\boldsymbol{c}\in A$，使$[\boldsymbol{c},\boldsymbol{x})\subset A$. 设$U$是$\boldsymbol{x}$的任一邻域. 显然$U\cap[\boldsymbol{c},\boldsymbol{x})\neq\Phi$，因而$U\cap A\neq\Phi$. 所以$\boldsymbol{x}\in\overline{A}$. 于是得$A^a\subset\overline{A}$.

定义 1.4.2 设$C\subset E$为凸集，若$\text{int}(C)\neq\Phi$，则称C为**凸体**(convex body).

由定理1.4.4知，若C为凸体，则C也是凸代数体；并且还可以证明以下定理.

定理 1.4.5 设$C\subset E$是凸体，则

1° $\operatorname{int}(C)=\operatorname{int}(\overline{C})=C^i$

2° $\overline{C}=\overline{\operatorname{int}(C)}=C^a$.

证 1° 设$\boldsymbol{x}\in C^i$，$\boldsymbol{v}\in\operatorname{int}(C)$. 于是存在$\boldsymbol{z}$使$\boldsymbol{x}\in(\boldsymbol{z},\boldsymbol{v})$且$(\boldsymbol{z},\boldsymbol{v})\subset C$，因而$\boldsymbol{z}\in\overline{C}$. 依定理1.4.1$(\boldsymbol{z},\boldsymbol{v}]\subset\operatorname{int}(C)$. 于是$\boldsymbol{x}\in\operatorname{int}(C)$. 由此得$C^i\subset\operatorname{int}(C)$. 再考虑到定理1.4.4即得$\operatorname{int}(C)=C^i$.

再设$\boldsymbol{y}\in\operatorname{int}(\overline{C})$，$\boldsymbol{v}\in\operatorname{int}(C)$. 于是存在$\boldsymbol{z}$使$\boldsymbol{y}\in(\boldsymbol{z},\boldsymbol{v})$，且$[\boldsymbol{z},\boldsymbol{y}]\subset\overline{C}$. 由定理1.4.1得$(\boldsymbol{z},\boldsymbol{v}]\subset\operatorname{int}(C)$，所以$\boldsymbol{y}\in\operatorname{int}(C)$. 于是

$$\operatorname{int}(\overline{C})\subset\operatorname{int}(C).$$

而$\operatorname{int}(C)\subset\operatorname{int}(\overline{C})$，所以$\operatorname{int}(C)=\operatorname{int}(\overline{C})$.

2° 设$\boldsymbol{x}\in\overline{C}$，$\boldsymbol{v}\in\operatorname{int}(C)$. 依定理1.4.1$[\boldsymbol{v},\boldsymbol{x})\subset\operatorname{int}(C)$，因而$[\boldsymbol{v},\boldsymbol{x})\subset C$. 依定义1.3.2 $\boldsymbol{x}\in C^a$. 于是$\overline{C}\subset C^a$. 再由定理1.4.4得$C^a\subset\overline{C}$，所以$\overline{C}=C^a$.

再由$[\boldsymbol{v},\boldsymbol{x})\subset\operatorname{int}(C)$得$\boldsymbol{x}\in\overline{\operatorname{int}(C)}$. 于是$\overline{C}\subset\overline{\operatorname{int}(C)}$. 而$\overline{\operatorname{int}(C)}\subset\overline{C}$，所以$\overline{C}=\overline{\operatorname{int}(C)}$.

§1.5 超 平 面

超平面是一类重要的仿射集.

设L为V的一个线性子空间. 若存在$\boldsymbol{d}\in V$，$\boldsymbol{d}\notin L$使

$$V=L+\mathscr{R}\boldsymbol{d}=\{\boldsymbol{m}+\lambda\boldsymbol{d}\mid\boldsymbol{m}\in L,\ \lambda\in\mathscr{R}\},$$

则称L具有**余维数**（co-dimension）1.

定义 1.5.1 设L为V中具有余维数1的线性子空间，$\boldsymbol{a}\in V$. 则称仿射集$H=L+\boldsymbol{a}$为V的**超平面**（hyperplane）.

定理 1.5.1 H为V中的一张超平面的充分必要条件是：存在一个非零实线性函数f：$V\to\mathscr{R}$与$a\in\mathscr{R}$使$H=f^{-1}(a)$（即$H=\{\boldsymbol{x}\mid f(\boldsymbol{x})=a\}$）.

证 充分性. 设存在一个非零实线性函数$f: V \to \mathscr{R}$与$\alpha \in \mathscr{R}$使$H = f^{-1}(\alpha)$.

因为f是一个实线性函数，所以$f^{-1}(0) = \{\boldsymbol{x} \mid f(\boldsymbol{x}) = 0\}$是一个线性子空间. 因为$f$非零，所以存在$\boldsymbol{d} \in V$，使$f(\boldsymbol{d}) = 1$. $\forall \boldsymbol{x} \in V$有$f(\boldsymbol{x} - f(\boldsymbol{x})\boldsymbol{d}) = f(\boldsymbol{x}) - f(\boldsymbol{x}) = 0$，因此 $\boldsymbol{x} - f(\boldsymbol{x})\boldsymbol{d} \in f^{-1}(0)$，从而$\boldsymbol{x} \in f^{-1}(0) + f(\boldsymbol{x})\boldsymbol{d}$. 于是$V \subset f^{-1}(0) + \mathscr{R}\boldsymbol{d}$. 而$f^{-1}(0) + \mathscr{R}\boldsymbol{d} \subset V$，所以

$$f^{-1}(0) + \mathscr{R}\boldsymbol{d} = V,$$

即$f^{-1}(0)$为具有余维数1的线性子空间.

再由$H = f^{-1}(\alpha) = \alpha\boldsymbol{d} + f^{-1}(0)$得$H = f^{-1}(\alpha)$是一张超平面.

必要性. 设$H = L + \boldsymbol{a}$是V中一张超平面，其中$\boldsymbol{a} \in V$，L为余维数为1的线性子空间. 取$\boldsymbol{d} \in V, \boldsymbol{d} \overline{\in} L$. 对于$V$中每个$\boldsymbol{x}$均可分解为$\boldsymbol{m} + \lambda\boldsymbol{d}$，其中$\boldsymbol{m} \in L$，$\lambda \in \mathscr{R}$. 令$f(\boldsymbol{x}) = \lambda$. $f(\boldsymbol{x})$是一个非零实线性函数，而且$f(\boldsymbol{d}) = 1$. 当$\boldsymbol{x} \in L$时$f(\boldsymbol{x}) = 0$. 再令$\alpha = f(\boldsymbol{a})$. 显然有

$$H = L + \boldsymbol{a} \subset \{\boldsymbol{x} \mid f(\boldsymbol{x}) = \alpha\} = f^{-1}(\alpha).$$

还可以证明$f^{-1}(\alpha) \subset H$，从而$H = f^{-1}(\alpha)$.

在$\mathscr{R}^n$中余维数为1的线性子空间是$\{\boldsymbol{x} \mid \boldsymbol{p}^T\boldsymbol{x} = 0\}$，其中$\boldsymbol{p} = (p_1, \cdots, p_n)^T \neq 0$. 超平面是$n-1$维仿射集$H = \{\boldsymbol{x} \mid \boldsymbol{p}^T\boldsymbol{x} = \alpha\}$，其中$\boldsymbol{p} = (p_1, \cdots, p_n)^T \neq 0$，$\alpha \in \mathscr{R}$. 称$\boldsymbol{p}$为超平面的法矢. $\mathscr{R}^n$中的超平面H简记为$\boldsymbol{p}^T\boldsymbol{x} = \alpha$.

定理 1.5.2 设C，$D \subset V$是非空凸集，$C \cap D = \Phi, C \cup D = V$，$H = C^a \cap D^a$. 若$H \neq V$，则$H$是一张超平面.

证 设$\boldsymbol{a} \in V \setminus H$，$\boldsymbol{h} \in H$. 因为$H = (H - \boldsymbol{h}) + \boldsymbol{h}$，所以要证明$H$是一张超平面，只要证明线性子空间$H - \boldsymbol{h}$具有余维数1就够了（由定理1.3.3可知$H$是一仿射集）.

因为$\boldsymbol{a} \overline{\in} H$，由定理1.3.3可知，若$\boldsymbol{a} \overline{\in} D^a$，则$\boldsymbol{a} \in C^i$；若$\boldsymbol{a} \overline{\in} C^a$，则$\boldsymbol{a} \in D^i$，所以$\boldsymbol{a} \in C^i \cup D^i$. 不妨设$\boldsymbol{a} \in C^i$，于是$2\boldsymbol{h} - \boldsymbol{a} \overline{\in} H$

(否则，将有$\boldsymbol{a}\in H$，与假设矛盾)．从而$2\boldsymbol{h}-\boldsymbol{a}\in C^i\cup D^i$．用反证法可得$2\boldsymbol{h}-\boldsymbol{a}\overline{\in}C^i$，因而$2\boldsymbol{h}-\boldsymbol{a}\in D^i$．

任取$\boldsymbol{x}\in V$．若$\boldsymbol{x}+\boldsymbol{h}\in C$，则$[\boldsymbol{x}+\boldsymbol{h}, 2\boldsymbol{h}-\boldsymbol{a})\cap H\neq\Phi$(依定理1.3.3)．所以存在$\boldsymbol{v}\in H$与$\lambda\in(0,1]$使

$$\boldsymbol{v}=\lambda(\boldsymbol{x}+\boldsymbol{h})+(1-\lambda)(2\boldsymbol{h}-\boldsymbol{a}),$$

由此得

$$\boldsymbol{x}=\left(\frac{1}{\lambda}-1\right)(\boldsymbol{a}-\boldsymbol{h})+\frac{1}{\lambda}(\boldsymbol{v}-\boldsymbol{h}).$$

若 $\boldsymbol{x}+\boldsymbol{h}\in D$，则$[\boldsymbol{x}+\boldsymbol{h}, \boldsymbol{a})\cap H\neq\Phi$．所以存在$\boldsymbol{v}\in H$与$\lambda\in(0, 1]$使

$$\boldsymbol{v}=\lambda(\boldsymbol{x}+\boldsymbol{h})+(1-\lambda)\boldsymbol{a}.$$

由此得

$$\boldsymbol{x}=\left(1-\frac{1}{\lambda}\right)(\boldsymbol{a}-\boldsymbol{h})+\frac{1}{\lambda}(\boldsymbol{v}-\boldsymbol{h}).$$

综上所述，$\forall\,\boldsymbol{x}\in V$都存在$\boldsymbol{y}\in H-\boldsymbol{h}$，$\mu\in\mathscr{R}$，使

$$\boldsymbol{x}=\boldsymbol{y}+\mu(\boldsymbol{a}-\boldsymbol{h})\subset H-\boldsymbol{h}+\mathscr{R}(\boldsymbol{a}-\boldsymbol{h}).$$

由此可得

$$V=H-\boldsymbol{h}+\mathscr{R}(\boldsymbol{a}-\boldsymbol{h}).$$

这表明线性子空间$H-\boldsymbol{h}$具有余维数1．因而，H是一张超平面．

定理 1.5.3 设A为$m\times n$矩阵，$\boldsymbol{b}\in\mathscr{R}^m$．则$M=\{\boldsymbol{x}\mid A\boldsymbol{x}=\boldsymbol{b}\}$是仿射集．

证 设A的第i行为$\boldsymbol{a}^{(i)T}$，则$\boldsymbol{a}^{(i)T}\boldsymbol{x}=b_i$为一张超平面，记为$H_i$．每个超平面$H_i$都是仿射集，因而它们的交$M=\bigcap_{i=1}^{m}H_i=\{\boldsymbol{x}\mid A\boldsymbol{x}=\boldsymbol{b}\}$也是仿射集(根据定理1.2.1的推论)．

***定理 1.5.4** 设$H=f^{-1}(a)$为E中一张超平面，$D=\{\boldsymbol{x}\mid f(\boldsymbol{x})>a\}$，其中$f$: $E\to\mathscr{R}$为线性函数，则下列条件等价[2]：

1° H是闭的；

2° $\text{int}(D)\neq\Phi$；

3° f 是连续的.

（证明从略）

有限维线性空间$\mathscr{R}^n$中的每个线性函数 $f(\boldsymbol{x})$（即 $\boldsymbol{p}^T\boldsymbol{x}$，其中 $\boldsymbol{p}\in\mathscr{R}^n$）都是连续的，因而$\mathscr{R}^n$中的超平面$H=\{\boldsymbol{x}\,|\,\boldsymbol{p}^T\boldsymbol{x}=\alpha\}$都是闭的．在无限维线性空间中存在不连续的线性函数[2].

定义 1.5.2 设f：$E\to\mathscr{R}$为连续线性函数，则称$\{\boldsymbol{x}\,|\,f(\boldsymbol{x})\leqslant\alpha\}$与$\{\boldsymbol{x}\,|\,f(\boldsymbol{x})\geqslant\alpha\}$为闭半空间，$\{\boldsymbol{x}\,|\,f(\boldsymbol{x})<\alpha\}$与$\{\boldsymbol{x}\,|\,f(\boldsymbol{x})>\alpha\}$为开半空间.

定义 1.5.3 设E为赋范线性空间，则E上所有连续线性函数构成的线性空间E'称为E的对偶.

$\mathscr{R}^n$上的连续性函数$u(\boldsymbol{x})=\boldsymbol{u}^T\boldsymbol{x}$，其中$\boldsymbol{u}\in\mathscr{R}^n$，与$\boldsymbol{u}$对应，所以$\mathscr{R}^n$的对偶还是$\mathscr{R}^n$．设$\boldsymbol{x}\in E$，$\boldsymbol{u}\in E'$．为方便起见，将$\boldsymbol{u}(\boldsymbol{x})$写成$(\boldsymbol{x}\,|\,\boldsymbol{u})$．设$A\subset E$．有时用$(A\,|\,\boldsymbol{u})\geqslant\alpha$表示$(\boldsymbol{x}\,|\,\boldsymbol{u})\geqslant\alpha$，$\forall\,\boldsymbol{x}\in A$.

若$\boldsymbol{u}\in E'$，$\alpha\in\mathscr{R}$，则

$$F(\boldsymbol{x},\ \lambda)=(\boldsymbol{x}\,|\,\boldsymbol{u})+\alpha\lambda,\ (\boldsymbol{x},\ \lambda)\in E\times\mathscr{R}$$

表示$E\times\mathscr{R}$上的连续线性函数．记作$F=(\boldsymbol{u},\alpha)$.

$E\times\mathscr{R}$中的闭超平面H可以写成

$$(\boldsymbol{x}\,|\,\boldsymbol{u})+\alpha\lambda=\beta,\ \boldsymbol{x}\in E,\ \lambda\in\mathscr{R}. \tag{1.5-1}$$

其中$(\boldsymbol{u},\alpha)\neq(0,0)$，$\beta\in\mathscr{R}$．称$(\boldsymbol{u},\alpha)$为超平面$H$的法矢．若$\alpha=0$，则称$H$为垂直的超平面．若$\alpha\neq0$，则称$H$为非垂直的超平面.非垂直的闭超平面可以写成

$$\lambda=\left(\boldsymbol{x}\,\middle|\,-\frac{1}{\alpha}\boldsymbol{u}\right)+\frac{\beta}{\alpha},\ \boldsymbol{x}\in E. \tag{1.5-2}$$

§1.6 凸锥与多面锥

本节将讨论凸锥与多面锥.

1.6.1 凸锥

定义 1.6.1 设$K\subset V$. 若$\forall\, \boldsymbol{x}\in K$, $\forall\, \lambda>0$ 均有$\lambda\boldsymbol{x}\in K$, 则称$K$为锥.

设$K\subset V$. 若K是锥，而且是凸集，则称K为 **凸锥**（convex cone）.

点0为锥的顶点，它可以属于锥K，也可以不属于锥K. 若$K\neq\Phi$,则点$0\in\overline{K}$.

定理 1.6.1 设$K\subset V$. K为凸锥的充分必要条件是：$\forall\,\boldsymbol{x}, \boldsymbol{y}\in K$, $\forall\,\lambda>0$，均有

$$\boldsymbol{x}+\boldsymbol{y}\in K,\ \lambda\boldsymbol{x}\in K.$$

证 必要性. 设K为凸锥. $\boldsymbol{x},\boldsymbol{y}\in K$, $\lambda>0$. 由锥的定义得$\lambda\boldsymbol{x}\in K$. 再由$K$的凸性得$\frac{1}{2}\boldsymbol{x}+\frac{1}{2}\boldsymbol{y}\in K$. 于是 $\boldsymbol{x}+\boldsymbol{y}=2\left(\frac{1}{2}\boldsymbol{x}+\frac{1}{2}\boldsymbol{y}\right)\in K$.

充分性. 设$\forall\,\boldsymbol{x},\boldsymbol{y}\in K$, $\forall\lambda>0$ 均有

$$\boldsymbol{x}+\boldsymbol{y}\in K,\ \lambda\boldsymbol{x}\in K.$$

由定义可知K是锥. 再设$\mu\in(0,1)$, 于是$\mu\boldsymbol{x}\in K$, $(1-\mu)\boldsymbol{y}\in K$. 相加得$\mu\boldsymbol{x}+(1-\mu)\boldsymbol{y}\in K$. 因而$K$是凸集，所以$K$是凸锥.

推论 设K_1, $K_2\subset V$均为凸锥，则和集K_1+K_2是凸锥，交集$K_1\cap K_2$也是凸锥.

定义 1.6.2 设$K\subset E$是一个锥，

$$K^0=\{\boldsymbol{u}\,|\,\boldsymbol{u}\in E',\ (K\,|\,\boldsymbol{u})\leqslant 0\},$$

$$K^{00}=\{\boldsymbol{x}\,|\,\boldsymbol{x}\in E,\ (\boldsymbol{x}\,|\,K^0)\leqslant 0\},$$

则称K^0为锥K的**极**（polar），K^{00}为锥K的**双极**（bipolar）.

其中记号$(K\,|\,\boldsymbol{u})\leqslant 0$表示$\forall\,\boldsymbol{x}\in K$均有$(\boldsymbol{x}\,|\,\boldsymbol{u})\leqslant 0$. 记号$(\boldsymbol{x}\,|\,K^0)\leqslant 0$表示$\forall\,\boldsymbol{u}\in K^0$均有$(\boldsymbol{x}\,|\,\boldsymbol{u})\leqslant 0$.

定理 1.6.2 设$K\subset E$是锥,则K的极K^0与K的双极K^{00}都是包含点0的凸锥.

证 显然$0\in K^0$, $0\in K^{00}$.下面证明K^0为凸锥.设$\boldsymbol{u},\boldsymbol{v}\in K^0$,

$\lambda > 0$，则

$$(K|\lambda \boldsymbol{u}) = \lambda (K|\boldsymbol{u}) \leqslant 0,$$

$$(K|\boldsymbol{u}+\boldsymbol{v}) = (K|\boldsymbol{u}) + (K|\boldsymbol{v}) \leqslant 0,$$

因而 $\lambda \boldsymbol{u}, \boldsymbol{u}+\boldsymbol{v} \in K^0$. 依定理1.6.1 K^0是凸锥.

同理可证K^{00}是凸锥.

定理 1.6.3 设K为赋范线性空间的非空开凸锥，则$K+\overline{K}=K$.

证 因为 $0\in\overline{K}$，$K+0=K$，所以$K\subset K+\overline{K}$.

因为K是开集，所以$K=\text{int}(K)$. 设$\boldsymbol{x}\in K$，$\boldsymbol{y}\in\overline{K}$. 依定理1.4.1得$(\boldsymbol{x},\boldsymbol{y})\subset\text{int}(K)$，因而$\frac{1}{2}(\boldsymbol{x}+\boldsymbol{y})\in\text{int}(K)=K$. 因为$K$是锥，所以

$$\boldsymbol{x}+\boldsymbol{y}=2\left(\frac{1}{2}(\boldsymbol{x}+\boldsymbol{y})\right)\in K.$$

由$\boldsymbol{x},\boldsymbol{y}$的任意性得$K+\overline{K}\subset K$.

综上所述得 $K+\overline{K}=K$.

定理 1.6.4 设K_1, K_2是赋范线性空间中的两个非空的锥，则

1° $(K_1+K_2)^0=K_1^0\cap K_2^0=(K_1\cup K_2)^0$;

2° $(K_1\cap K_2)^0\supset K_1^0+K_2^0=\text{co}(K_1^0\cup K_2^0)$.

（证明留给读者）

1.6.2 多面锥

定义 1.6.3 $\mathscr{R}^n$中有限多个包含边界点0的闭半空间的交称为$\mathscr{R}^n$的一个**多面锥**（polyhedral cone）.

设K为$\mathscr{R}^n$的一个多面锥，则它可以表示为

$$K=\{\boldsymbol{x}\,|\,\boldsymbol{p}^{(i)T}\boldsymbol{x}\leqslant 0,\ i=1,\cdots,m,\ \boldsymbol{x}\in\mathscr{R}^n\}$$

其中$\boldsymbol{p}^{(i)}\in\mathscr{R}^n$，$\boldsymbol{p}^i\neq 0$.

设P^m为$\mathscr{R}^m$中的非负卦限$\{\boldsymbol{y}\,|\,\boldsymbol{y}\geqslant 0,\ \boldsymbol{y}\in\mathscr{R}^m\}$，$T$为由$\mathscr{R}^n$到

$\mathscr{R}^m$的线性映射：

$$(Tx)_i = p^{(i)T}x, \quad x, \ p^{(i)} \in \mathscr{R}^n, \quad i = 1, \cdots, m,$$

可以将多面锥$K \subset \mathscr{R}^n$表示为

$$K = \{x \mid Tx \leqslant 0\} = -T^{-1}P^m.$$

因为包含边界点0的闭半空间是包含点0的凸锥，而有限多个包含点0的凸锥的交仍然是包含点0的凸锥，所以多面锥是包含点 0 的凸锥.

定义1.6.4 设$a^{(i)} \in \mathscr{R}^n$，$a^{(i)} \neq 0$，$i = 1, \cdots, m$，而

$$K = \left\{ x \,\middle|\, x = \sum_{i=1}^{m} \lambda_i a^{(i)}, \ \forall \lambda_1, \cdots, \lambda_m \geqslant 0 \right\},$$

则称K为有限生成的凸锥，称$a^{(1)}, \cdots, a^{(m)}$为K的生成元.

设Q为由$\mathscr{R}^m$到$\mathscr{R}^n$的线性映射，则集合

$$K = \{x \mid x = Qy, \ y \geqslant 0, \ y \in \mathscr{R}^m\} = QP^m$$

是$\mathscr{R}^n$中有限生成的凸锥. $\mathscr{R}^n$中有限生成的凸锥都可以写成 这种形式.

定理 1.6.5 设K为$\mathscr{R}^n$中有限生成的凸锥. 则K的极K^0是一个多面锥.

证 设$\mathscr{R}^n$中有限生成的凸锥$K = QP^m$，其中Q 为 由$\mathscr{R}^m$到$\mathscr{R}^n$的线性映射，P^m为$\mathscr{R}^m$中的非负卦限，则K的极为

$$K^0 = \{u \mid (K \mid u) \leqslant 0, \ u \in \mathscr{R}^n\}.$$

因为 $(K \mid u) \leqslant 0$ 即 $(QP^m \mid u) \leqslant 0$. 它等价于

$$(Qt \mid u) \leqslant 0, \quad \forall t \in P^m,$$

即$(Qt)^T u = t^T Q^T u \leqslant 0$，$\forall t \in P^m$. 这又等价于$Q^T u \leqslant 0$，于是得$K^0 = \{u \mid Q^T u \leqslant 0, \ u \in \mathscr{R}^n\}$，即

$$K^0 = \{x \mid Q^T x \leqslant 0, \ x \in \mathscr{R}^n\},$$

其中Q^T是由$\mathscr{R}^n$到$\mathscr{R}^m$的线性映射. 因而K^0是一个多面锥.

定理 1.6.6 设K为$\mathscr{R}^n$中有限生成的凸锥，则

1° K是有限多个有限生成的凸锥的并，而其中每个凸锥都

具有一组线性无关的生成元；

2° K是闭的.

证 依定义1.6.4存在$\boldsymbol{\alpha}^{(1)}, \cdots, \boldsymbol{\alpha}^{(m)} \in \mathscr{R}^n$，使

$$K=\{\lambda_1\boldsymbol{\alpha}^{(1)}+\cdots+\lambda_m\boldsymbol{\alpha}^{(m)} \mid \lambda_1, \cdots, \lambda_m \geqslant 0\}. \quad (1.6\text{-}1)$$

1° 设$\boldsymbol{x}\in K$，则

$$\boldsymbol{x}=\lambda_1\boldsymbol{\alpha}^{(1)}+\cdots+\lambda_m\boldsymbol{\alpha}^{(m)}, \ \lambda_1, \cdots, \lambda_m \geqslant 0.$$

若$\boldsymbol{\alpha}^{(1)}, \cdots, \boldsymbol{\alpha}^{(m)}$线性相关，则存在不全为零的数$\mu_1, \cdots, \mu_m$，使

$$\mu_1\boldsymbol{\alpha}^{(1)}+\cdots+\mu_m\boldsymbol{\alpha}^{(m)}=0.$$

不妨设其中至少有一个$\mu_i>0$，于是得

$$\boldsymbol{x}=\sum_{i=1}^{m}(\lambda_i-\rho\mu_i)\boldsymbol{\alpha}^{(i)}, \ \forall \rho\in\mathscr{R}. \quad (1.6\text{-}2)$$

令

$$\rho=\min_{1\leq i\leq m}\left\{\frac{\lambda_i}{\mu_i} \,\middle|\, \mu_i>0\right\},$$

则式(1.6-2)中所有系数均非负，而且至少有一个为零。于是$\boldsymbol{x}$表为$\{\boldsymbol{\alpha}^{(1)}, \cdots, \boldsymbol{\alpha}^{(m)}\}$中至多$m-1$个元素的非负组合. 若它们仍然线性相关，则重复以上步骤，直至所余元素线性无关为止。于是$\boldsymbol{x}$属于由$\{\boldsymbol{\alpha}^{(1)}, \cdots, \boldsymbol{\alpha}^{(m)}\}$的一个具有线性无关元素的子集所生成的凸锥。因为$\{\boldsymbol{\alpha}^{(1)}, \cdots, \boldsymbol{\alpha}^{(m)}\}$只有有限多个具有线性无关元素的子集，所以命题得证.

§1.7 凸多胞形 · 极点 · 极方向

本节引入凸多胞形、极点与极方向的概念. 极点是凸集的一个重要概念，极方向是无界凸集的一个重要概念.

1.7.1 凸多胞形

定义 1.7.1 E中有限多个闭半空间的交称为E的一个**凸多胞形**（convex polytope）.

显然，多面锥是凸多胞形，而凸多胞形不一定是多面锥（因为凸多胞形不一定含点0）.

设S是$\mathscr{R}^n$中的一个凸多胞形，则它可以表示为

$$S=\{\boldsymbol{x}\mid \boldsymbol{p}^{(i)T}\boldsymbol{x}\leqslant b_i,\quad i=1,\cdots,m\},$$

其中$\boldsymbol{p}^{(i)}\in\mathscr{R}^n$，$\boldsymbol{p}^{(i)}\neq 0$，$b_i\in\mathscr{R}$。容易看出，$S$也可表示为

$$S=\{\boldsymbol{x}\mid \boldsymbol{p}^{(i)T}\boldsymbol{x}\geqslant b_i,\quad i=1,\cdots,m\}.$$

因为m个线性等式

$$\boldsymbol{p}^{(i)T}\boldsymbol{x}=b_i,\quad i=1,\cdots,m$$

可以化为$m+1$个线性不等式

$$\boldsymbol{q}^{(i)T}\boldsymbol{x}\leqslant b_i,\quad i=1,\cdots,m+1,$$

其中

$$\boldsymbol{p}^{(m+1)}=-\sum_{i=1}^{m}\boldsymbol{p}^{(i)},\quad b_{m+1}=-\sum_{i=1}^{m}b_i,$$

所以

$$\{\boldsymbol{x}\mid \boldsymbol{p}^{(i)T}\boldsymbol{x}=b_i,\quad i=1,\cdots,m\}$$

也表示$\mathscr{R}^n$中的凸多胞形.

1.7.2 极点

定义 1.7.2 设$C\subset V$为凸集.若C中不存在两个相异的点$\boldsymbol{x}^{(1)}$，$\boldsymbol{x}^{(2)}$及实数$\alpha\in(0,1)$，使

$$\boldsymbol{x}=\alpha\boldsymbol{x}^{(1)}+(1-\alpha)\boldsymbol{x}^{(2)},$$

则称$\boldsymbol{x}$为C的**极点**（extreme point）.

例1 设$C=\{\boldsymbol{x}\mid\|\boldsymbol{x}\|\leqslant r\}\subset\mathscr{R}^2$.试证圆周$\|\boldsymbol{x}\|=r$上的点都是它的极点.

证 由于对称性只需证明圆周上的一个点，如$\boldsymbol{x}^{(0)}=(r,0)^T$，是极点．设存在$\boldsymbol{x}^{(1)}$，$\boldsymbol{x}^{(2)}\in C$，$\alpha\in(0,1)$，使

$$\alpha\boldsymbol{x}^{(1)}+(1-\alpha)\boldsymbol{x}^{(2)}=(r,0)^T,$$

即

$$\alpha x_1^{(1)}+(1-\alpha)x_1^{(2)}=r,$$

$$\alpha x_2^{(1)}+(1-\alpha)x_2^{(2)}=0.$$

因为$\boldsymbol{x}^{(1)}$，$\boldsymbol{x}^{(2)}\in C$，因而$x_1^{(1)}$，$x_1^{(2)}\leqslant r$．于是$x_1^{(1)}=x_1^{(2)}=r$，$x_2^{(1)}=x_2^{(2)}=0$，即$\boldsymbol{x}^{(1)}=\boldsymbol{x}^{(2)}=\boldsymbol{x}^{(0)}$所以$\boldsymbol{x}^{(0)}$是极点．

例 2 设$C=co(\boldsymbol{x}^{(1)},\cdots,\boldsymbol{x}^{(k)})$，$\boldsymbol{x}^{(1)}$，$\cdots$，$\boldsymbol{x}^{(k)}$为$C$的顶点，则$x^{(1)}$，$\cdots$，$x^{(k)}$都是$C$的极点．

（证明留给读者）

顺便指出，并非每个凸集都有极点．显然，$\mathscr{R}^n$无极点，超球$\{\boldsymbol{x}|\ \|\boldsymbol{x}\|<r\}$也无极点．

定理 1.7.1 设凸多胞形为

$$S=\{\boldsymbol{x}\ |A\boldsymbol{x}=\boldsymbol{b},\ \boldsymbol{x}\geqslant 0\} \tag{1.7-1}$$

其中A为$m\times n$矩阵，秩为m，$\boldsymbol{b}\in\mathscr{R}^m$，则$\boldsymbol{x}$为$S$的极点的充分必要条件是：$A=(B,\ N)$，其中$B$为$m\times m$的非奇异矩阵，而且

$$\boldsymbol{x}=\begin{pmatrix}\boldsymbol{x}_B\\ \boldsymbol{x}_N\end{pmatrix}=\begin{pmatrix}B^{-1}\boldsymbol{b}\\ 0\end{pmatrix}\geqslant 0$$ [3]

推论 由式（1.7-1）确定的凸多胞形只能具有有限多个极点（不超过C_n^m个）．

定理 1.7.2 若由式（1.7-1）确定的凸多胞形$S\neq\Phi$，则S至少有一个极点[3]．

（证明从略）

1.7.3 极方向

定义 1.7.3 设$C\subset V$为凸集，$\boldsymbol{d}\in V$，$\boldsymbol{d}\neq 0$．若对于每一点$\boldsymbol{x}\in C$都有

$$\boldsymbol{x}+\lambda\boldsymbol{d}\in C,\quad \forall\lambda\geqslant 0$$

则称$\boldsymbol{d}$为C的一个**方向**．

设$\boldsymbol{d}^{(1)}$，$\boldsymbol{d}^{(2)}$是凸集C的方向．若存在$\alpha>0$，使$\boldsymbol{d}^{(2)}=\alpha\boldsymbol{d}^{(1)}$，则$\boldsymbol{d}^{(2)}$与$\boldsymbol{d}^{(1)}$是同一方向．否则，是不同方向．

定义 1.7.4 设$C\subset V$为凸集，$\boldsymbol{d}$为C的一个方向．若不存在两个不同方向$\boldsymbol{d}^{(1)}$，$\boldsymbol{d}^{(2)}$与两个实数α_1，$\alpha_2>0$，使

$$\boldsymbol{d}=\alpha_1\boldsymbol{d}^{(1)}+\alpha_2\boldsymbol{d}^{(2)}$$

则称$\boldsymbol{d}$为C的**极方向**（extreme direction）．

由定义可知，有界集无方向，也无极方向．

例 3 设$K=\{\boldsymbol{x}\,|\,\boldsymbol{x}\geqslant 0\}\subset\mathscr{R}^2$．试证明$(1,0)^T$，$(0,1)^T$为$K$的极方向．

证 1° 因为点$\boldsymbol{0}=(0,0)^T\in K$，所以若$\boldsymbol{d}$为$K$的方向，则$\boldsymbol{d}\geqslant 0$，$\boldsymbol{d}\neq 0$．

2° 若$\boldsymbol{d}^{(1)}$，$\boldsymbol{d}^{(2)}$为K的方向，α_1，$\alpha_2>0$，使

$$\alpha_1\boldsymbol{d}^{(1)}+\alpha_2\boldsymbol{d}^{(2)}=(1,0)^T$$

则$d_2^{(1)}=d_2^{(2)}=0$（因为$d_2^{(1)}$，$d_2^{(2)}\geqslant 0$，α_1，$\alpha_2>0$）．因而$d_1^{(1)}$，$d_1^{(2)}>0$，所以存在$\alpha>0$，使$d_1^{(2)}=\alpha d_1^{(1)}$．于是$\boldsymbol{d}^{(2)}=\alpha\boldsymbol{d}^{(1)}$，即$\boldsymbol{d}^{(2)}$与$\boldsymbol{d}^{(1)}$是同一方向，故$(1,0)^T$为$K$的极方向．

同理可证$(0,1)^T$是K的极方向．

（容易证明，K的其它方向都不是极方向）

定理 1.7.3 设S由式(1.7-1)给出．则$\boldsymbol{d}$为S的方向的充分必要条件是：$A\boldsymbol{d}=0$，$\boldsymbol{d}\geqslant 0$，$\boldsymbol{d}\neq 0$．

（证明留给读者）

定理 1.7.4 设S由式(1.7-1)给出．若$A=(B,N)$，而且对于N的某一列$\boldsymbol{a}^{(j)}$有$B^{-1}\boldsymbol{a}^{(j)}\leqslant 0$，则向量

$$\boldsymbol{d}=\begin{pmatrix}-B^{-1}\boldsymbol{a}^{(j)}\\ \boldsymbol{e}^{(j)}\end{pmatrix}\tag{1.7-2}$$

是S的一个极方向，其中$\boldsymbol{e}^{(j)}\in\mathscr{R}^{n-m}$是一个单位列向量，对应于$\boldsymbol{a}^{(j)}$的分量为1，其余分量为零．

证 显然$\boldsymbol{d}\geqslant 0$，且$\boldsymbol{d}\neq 0$．因为$A\boldsymbol{d}=-\boldsymbol{a}^{(j)}+N\boldsymbol{e}^{(j)}=0$．所以由定理1.7.3知$\boldsymbol{d}$为$S$的一个方向．若$\boldsymbol{d}$不是极方向，则存在不同方向$\boldsymbol{d}^{(1)}\geqslant 0$，$\boldsymbol{d}^{(2)}\geqslant 0$与$\lambda_1$，$\lambda_2>0$，使

$$\boldsymbol{d}=\lambda_1\boldsymbol{d}^{(1)}+\lambda_2\boldsymbol{d}^{(2)} \tag{1.7-3}$$

因为$\boldsymbol{d}$有$n-m-1$个分量为零，由上式可知$\boldsymbol{d}^{(1)}$，$\boldsymbol{d}^{(2)}$相应的分量也都为零.

因为$\boldsymbol{d}$对应于$\boldsymbol{a}^{(j)}$的分量$d_j=1$，所以$\boldsymbol{d}^{(1)}$，$\boldsymbol{d}^{(2)}$对应的分量不能都是零. 不妨设$\boldsymbol{d}^{(1)}$对应的分量为1，即

$$\boldsymbol{d}^{(1)}=\begin{pmatrix}\boldsymbol{d}^{(11)}\\ \boldsymbol{e}^{(j)}\end{pmatrix}$$

因为$\boldsymbol{d}^{(1)}$是S的方向，所以$A\boldsymbol{d}^{(1)}=0$，即$B\boldsymbol{d}^{(11)}+N\boldsymbol{e}^{(j)}=0$. 由此可得$\boldsymbol{d}^{(11)}=-B^{-1}\boldsymbol{a}^{(j)}$，因而$\boldsymbol{d}^{(1)}=\boldsymbol{d}$.

若$\boldsymbol{d}^{(2)}$对应于$\boldsymbol{a}^{(j)}$的分量为零，则$\lambda_1=1$. 于是$\lambda_2=0$，与假设矛盾. 所以$\boldsymbol{d}^{(2)}$对应于$\boldsymbol{a}^{(j)}$的分量不为零，必定为正. 因而可设

$$\boldsymbol{d}^{(2)}=\alpha\begin{pmatrix}\boldsymbol{d}^{(21)}\\ \boldsymbol{e}^{(j)}\end{pmatrix},$$

其中$\alpha>0$. 因为$\boldsymbol{d}^{(2)}$是S的方向，所以$A\boldsymbol{d}^{(2)}=0$ 即 $B\boldsymbol{d}^{(21)}+N\boldsymbol{e}^{(j)}=0$. 由此得 $\boldsymbol{d}^{(21)}=-B^{-1}\boldsymbol{a}^{(j)}$，因而 $\boldsymbol{d}^{(2)}=\alpha\boldsymbol{d}^{(1)}$. 这表明$\boldsymbol{d}^{(2)}$与$\boldsymbol{d}^{(1)}$是相同方向. 所以$\boldsymbol{d}$是$S$的极方向.

例4 设S由方程组

$$\begin{cases}2x_1+3x_2-x_3-4x_4=8\\ -x_1+2x_2-6x_3+7x_4=3\end{cases}$$

的全部非负解$\boldsymbol{x}$确定. 试确定S的一个极方向.

解 设

$$B=\begin{pmatrix}3 & -1\\ 2 & -6\end{pmatrix},\quad N=\begin{pmatrix}2 & -4\\ -1 & 7\end{pmatrix}$$

于是

$$B^{-1}=\begin{pmatrix}\frac{6}{16} & \frac{-1}{16}\\ \frac{2}{16} & \frac{-3}{16}\end{pmatrix}$$

$a^{(4)}=(-4,\ 7)^T$，$B^{-1}a^{(4)}=\left(\frac{-31}{16},\ \frac{-29}{16}\right)^T\leqslant 0$．所以可得$S$的一个极方向

$$d=\left(0,\ \frac{31}{16},\ \frac{29}{16},\ 1\right)^T.$$

注 S的每个极方向都可以通过矩阵A的划分与$a^{(j)}$的选择（如定理1.7.4）得到．因而S只能具有有限多个（$\leqslant(n-m)C_n^m$）极方向．

§1.8 $\mathscr{R}^n$中的几个经典定理

本节将研究$\mathscr{R}^n$中的几个经典定理．

定理 1.8.1 （Caratheodory定理）设$A\subset\mathscr{R}^n$，$\dim(A)=k$．

1° $\forall x\in co(A)$，A中存在$m(\leqslant k+1)$个仿射无关点，使x为这m个点的凸组合；

2° $co(A)$等于顶点属于A的所有k维单纯形的并．[2]

证 设$x\in co(A)$．1° 根据定理 1.1.2 存在$x^{(1)},\cdots,x^{(p)}\in A$，使$x$为它们的凸组合，即

$$x=\sum_{i=1}^{p}\lambda_i x^{(i)},\tag{1.8-1}$$

其中 $\lambda_i\geqslant 0$，$\sum_{i=1}^{p}\lambda_i=1$．

若$p>k+1$，则$x^{(1)},\cdots,x^{(p)}$必仿射相关，即存在不全为零的数$\mu_1,\cdots,\mu_p$，使

$$\begin{cases}\mu_1x^{(1)}+\cdots+\mu_px^{(p)}=0,\\ \mu_1+\cdots+\mu_p\qquad=0,\end{cases}\tag{1.8-2}$$

其中至少有一个$\mu_i>0$. 令

$$\rho=\min_{1\leq i\leq p}\left\{\frac{\lambda_i}{\mu_i}\,\middle|\,\mu_i>0\right\}.$$

不妨设$\rho=\frac{\lambda_p}{\mu_p}$. 于是由式(1.8-1)与式(1.8-2)得

$$\boldsymbol{x}=\sum_{i=1}^{p-1}(\lambda_i-\rho\mu_i)\boldsymbol{x}^{(i)}=\sum_{i=1}^{p-1}\alpha_i\boldsymbol{x}^{(i)},$$

$$\alpha_i=\lambda_i-\rho\mu_i\geqslant 0,\quad \alpha_1+\cdots+\alpha_{p-1}=1.$$

若$p-1>k+1$, 则$\boldsymbol{x}^{(1)}, \cdots, \boldsymbol{x}^{(p-1)}$仍然仿射相关. 重复以上过程, 再减少一个点$\boldsymbol{x}^{(i)}$. 经有限步之后, 必可将$\boldsymbol{x}$表示为$A$中不多于$k+1$个点的凸组合.

若所得一组点仿射相关, 则还可重复以上过程, 直到最后所余的一组点仿射无关为止.

2° 若在1°中所得一组仿射无关点$\{\boldsymbol{x}^{(1)},\cdots,\boldsymbol{x}^{(m)}\}$的个数$m=k+1$, 则$\boldsymbol{x}$属于顶点在$A$中的一个$k$维单纯形. 若$m<k+1$, 则$A$中存在$k+1-m$个点$\{\boldsymbol{y}^{(m+1)},\cdots,\boldsymbol{y}^{(k+1)}\}$与它们构成具有$k+1$个点的仿射无关集$\{\boldsymbol{x}^{(1)}, \cdots, \boldsymbol{x}^{(m)}, \boldsymbol{y}^{(m+1)}, \cdots, \boldsymbol{y}^{(k+1)}\}$(见本章习题8). 于是

$$\boldsymbol{x}=\sum_{i=1}^{m}\alpha_i\boldsymbol{x}^{(i)}+0\cdot\boldsymbol{y}^{(m+1)}+\cdots+0\cdot\boldsymbol{y}^{(k+1)},$$

其中$\alpha_i\geqslant 0$, $\sum_{i=1}^{m}\alpha_i=1$. 上式表明

$$\boldsymbol{x}\in\mathrm{co}(\boldsymbol{x}^{(1)}, \cdots, \boldsymbol{x}^{(m)}, \boldsymbol{y}^{(m+1)}, \cdots, \boldsymbol{y}^{(k+1)}).$$

因而$\mathrm{co}(A)$含于顶点属于A的所有k维单纯形的并. 反包含关系是明显的. 于是命题得证.

推论1 设$\{C_i\,|\,i\in I\}$是$\mathscr{R}^n$中的一个凸集族, I为任意指标集, $C=\mathrm{co}(\bigcup_{i\in I}C_i)$, 则$\forall\boldsymbol{x}\in C$都可表示为属于不同$C_i$的最多$n+$

1个仿射无关点的凸组合.

证 设$A=\bigcup_{i\in I}C_i$，则$A\subset\mathscr{R}^n$，$\dim(A)=k\leqslant n$. 再设$\boldsymbol{x}\in C$依定理1.8.1 A中存在$m(m\leqslant k+1\leqslant n+1)$个仿射无关点$\boldsymbol{x}^{(1)}$，…，$\boldsymbol{x}^{(m)}$使

$$\boldsymbol{x}=\sum_{i=1}^{m}\lambda_i\boldsymbol{x}^{(i)},\ \lambda_i\geqslant 0,\quad \sum_{i=1}^{m}\lambda_i=1.$$

因为可以不考虑对应于$\lambda_i=0$的点$\boldsymbol{x}^{(i)}$，所以不妨设$\lambda_i>0$，$i=1$，…，m. 若$\boldsymbol{x}^{(1)}$，…，$\boldsymbol{x}^{(m)}$中有两个点，如$\boldsymbol{x}^{(1)}$，$\boldsymbol{x}^{(2)}$，属于同一个C_i，则令$\mu=\lambda_1+\lambda_2$，$\mu\boldsymbol{y}=\lambda_1\boldsymbol{x}^{(1)}+\lambda_2\boldsymbol{x}^{(2)}$，$\boldsymbol{y}\in C_i$
于是

$$\boldsymbol{x}=\mu\boldsymbol{y}+\sum_{i=3}^{m}\lambda_i\boldsymbol{x}^{(i)},$$

其中

$$\mu,\ \lambda_3,\ \cdots,\ \lambda_m>0,\ 且\mu+\lambda_3+\cdots+\lambda_m=1.$$

可以证明$\boldsymbol{y}$，$\boldsymbol{x}^{(3)}$，…，$\boldsymbol{x}^{(m)}$仍然仿射无关. 若其中还有两点属于同一个C_i，则依此类推可以再减少一个点，直到最后所得一组仿射无关点分别属于不同的C_i为止.

***推论2** 若$A\subset\mathscr{R}^n$是紧集，则co(A)也是紧集.

证 设$\dim(A)=k$. 依定理1.8.1 co(A)等于顶点属于A的所有k维单纯形的并，即由所有形如

$$\boldsymbol{y}=\sum_{i=1}^{k+1}\lambda_i\boldsymbol{x}^{(i)},\ \lambda_i\geqslant 0,\quad \sum_{i=1}^{k+1}\lambda_i=1$$

的点$\boldsymbol{y}$组成，其中$\boldsymbol{x}^{(i)}\in A$. 由

$$f(\boldsymbol{x}^{(1)},\ \cdots,\ \boldsymbol{x}^{(k+1)},\ \lambda_1,\ \cdots\lambda_{k+1})=\sum_{i=1}^{k+1}\lambda_i\boldsymbol{x}^{(i)}$$

所确定的映射$f:\mathscr{R}^{n(k+1)}\times\mathscr{R}^{k+1}\longrightarrow\mathscr{R}^n$是连续的，其中$\boldsymbol{x}^{(i)}\in\mathscr{R}^n$，

$\lambda_i \in \mathscr{R}$, $i=1$, …, $k+1$. 而A与集合

$$B=\left\{(\lambda_1, \cdots, \lambda_{k+1}) \middle| \lambda_i \geqslant 0, \quad \sum_{i=1}^{k+1} \lambda_i = 1\right\}$$

都是紧集，所以$A^{k+1}\times B$也是紧集. 因而作为$A^{k+1}\times B$的连续象集$f(A^{k+1}\times B)$的co(A)也是紧集.

定理 1.8.2 (Helly定理)[2,11] 设 $P=\{C_i \mid i\in I\}$是由$\mathscr{R}^n$中N($\geqslant n+1$)个非空凸集构成的集族. 若由其中任意$n+1$个凸集构成的子族的交非空，则整个集族$\{C_i\}$的交也非空，即$\bigcap_{i\in I} C_i \neq \Phi$.

证 设$I=\{1, \cdots, N\}$. 当$N=n+1$时，命题显然成立. 设当$N=k$ ($k\geqslant n+1$)时命题正确，考虑$N=k+1$时的情形. 对于每个$i\in\{1, \cdots, k+1\}$存在$\boldsymbol{x}^{(i)}\in\mathscr{R}^n$使

$$\boldsymbol{x}^{(i)} \in \bigcap_{j\neq i} C_j.$$

因为$k+1>n+1$，所以$\boldsymbol{x}^{(1)}$, …, $\boldsymbol{x}^{(k+1)}$仿射相关. 因而存在不全为零的数μ_1, …, $\mu_{k+1}\in\mathscr{R}$，使

$$\sum_{i=1}^{k+1} \mu_i \boldsymbol{x}^{(i)} = 0, \quad \sum_{i=1}^{k+1} \mu_i = 0. \tag{1.8-3}$$

由上式可知，μ_i中必有正的，也必有负的. 不妨设$\mu_1>0$, $\mu_2\geqslant 0$, …, $\mu_s\geqslant 0$, $\mu_{s+1}<0$, …, $\mu_{k+1}<0$. 设

$$\boldsymbol{y} = \sum_{i=1}^{s}\left(\mu_i \Big/ \sum_{i=1}^{s} \mu_i\right)\boldsymbol{x}^{(i)}, \tag{1.8-4}$$

则$\boldsymbol{y}$是$\boldsymbol{x}^{(1)}$, …, $\boldsymbol{x}^{(s)}$的凸组合. 而

$$\boldsymbol{x}^{(1)}, \cdots, \boldsymbol{x}^{(s)} \in \bigcap_{i=s+1}^{k+1} C_i,$$

所以
$$\boldsymbol{y} \in \bigcap_{i=s+1}^{k+1} C_i.$$

由式(1.8-3)与式(1.8-4)得

$$y = \left(-\sum_{i=s+1}^{k+1} \mu_i x^{(i)}\right) \Big/ \left(-\sum_{i=s+1}^{k+1} \mu_i\right)$$

$$= \sum_{i=s+1}^{k+1} \left[(-\mu_i) \Big/ \left(\sum_{i=s+1}^{k+1} (-\mu_i)\right)\right] x^{(i)},$$

因而y是$x^{(s+1)}$，…，$x^{(k+1)}$的凸组合．而$x^{(s+1)}$，…，$x^{(k+1)} \in \bigcap_{i=1}^{s} C_i$，所以 $y \in \bigcap_{i=1}^{s} C_i$，因而 $y \in \bigcap_{i=1}^{k+1} C_i$．由数学归纳法得

$$\bigcap_{i \in I} C_i \neq \Phi,$$

所以当N为任意自然数时命题都成立．

定义 1.8.1 设E为赋范线性空间，A，$B \subset E$．令

$$d(A, B) = \inf\{\|x-y\| \mid x \in A, y \in B\}, \tag{1.8-5}$$

则称$d(A, B)$为集合A与B之间的距离．

特别地，当$A=\{a\}$时，有a与B之间的距离

$$d(a, B) = \inf_{y \in B} \|a-y\|. \tag{1.8-6}$$

显然，若$a \in B$，则$d(a, B)=0$；若B为闭集，$a \overline{\in} B$，则$d(a, B)>0$．

***定义 1.8.2** 设$\mathscr{R}^n$为赋范线性空间．Ω为由$\mathscr{R}^n$中所有非空紧凸子集构成的集族，A，$B \in \Omega$．令

$$\alpha(A, B) = \sup\{d(x, B) \mid x \in A\}, \tag{1.8-7}$$

$$h(A, B) = \max\{\alpha(A, B), \alpha(B, A)\}, \tag{1.8-8}$$

称$h(A, B)$为A与B的Hausdorff距离．

设$r<0$，$\Omega(r)$为闭球$\{x \mid x \in \mathscr{R}^n, d(x, 0) \leqslant r\}$的所有

非空紧凸子集组成的集族，P为n维超立方体

$$P=\{x \mid |x_i|\leqslant r,\ \ i=1,\ \cdots,\ n\}.$$

*定义 1.8.3 将P的每条边等分为2^k段，相应地将P划分为棱长$s=r/2^{k-1}$的全等的闭超立方体．设P_k为这些闭超立方体的集合，$A\in\Omega(r)$．P_k中与A的交非空的所有闭超立方体的并称为A的一个k维覆盖．

*性质 若A，$B\in\Omega(r)$，且A，B有同样的k维覆盖，则$h(A,B)\leqslant\sqrt{n}\,r/2^{k-1}$．

*定义 1.8.4 设D_i，$A\in\mathscr{R}^n$，$i=1,2,\cdots$，均为非空凸集．若对于任给的正数ε都存在$N\in\mathscr{N}$，当$i>N$时不等式

$$h(D_i,\ A)\leqslant\varepsilon$$

恒成立，则称序列$\{D_i\}$收敛于A．记为$\lim\limits_{i\to\infty}D_i=A$．

$$(\lim_{i\to\infty}D_i=A \quad 即 \quad \lim_{i\to\infty}h(D_i,\ A)=0)$$

*定理 1.8.3 （Blaschke收敛定理）[2]具有Hausdoiff距离的集族$\Omega(r)$为紧度量空间．

证 因为具有Hausdorff距离的集族$\Omega(r)$是一个度量空间（见本章题17），所以只要证明$\Omega(r)$是列紧的即可．故只需证明：对于$\Omega(r)$的每个序列都有一个收敛的子列．

设$\{C_j\}$为$\Omega(r)$的一个序列．因为不同的1维覆盖的数目是有限的，所以$\{C_j\}$中必存在一个子序列$\{C_j^{(1)}\}$，它的元素具有同样的一维覆盖（否则，$\{C_j\}$为有限集）．同理，$\{C_j^{(1)}\}$中存在一个子序列$\{C_j^{(2)}\}$，它的元素具有同样的2维覆盖．依此类推得一系列序列$\{C_j^{(k)}\}$，$k=1,2,\cdots$它具有以下性质：

1° $\{C_j^{(k)}\}$的所有元素具有同一k维覆盖．

因此

$$h(C_i^{(k)},\ C_j^{(k)})\leqslant\sqrt{n}\ r/2^{k-1},\ \ \forall i,\ j\in\mathscr{N}.$$

2° $\{C_j^{(k)}\}$为$\{C_j^{(k-1)}\}$的子列．因此

$$h(C_i^{(p)},\ C_j^{(q)})\leqslant\sqrt{n}\ r/2^{q-1},\quad i,\ j,\ p,\ q\in\mathcal{N},\quad p>q$$

设 $\qquad D_i=C_i^{(i)},\quad i=1,\ 2,\ \cdots$

$$A_m=\overline{\mathrm{co}}\Big(\bigcup_{i\geqslant m}D_i\Big),\quad m=1,\ 2,\ \cdots$$

$$A=\bigcap_{m\geqslant 1}A_m.$$

因为A_m为非空紧凸集，A是A_1的闭子集，所以A也是非空紧凸集，从而$A\in\Omega(r)$.

$\{D_i\}\subset\{C_j\}$. 下面证明$\lim\limits_{i\to\infty}D_i=A$，即子列$\{D_i\}$收敛于$A$.

任给$\varepsilon>0$. 取$K\in\mathcal{N}$, $K\geqslant\log_2(2\sqrt{n}\,r/\varepsilon)$，则

$$h(D_j,\ D_m)\leqslant\varepsilon,\quad j,m>K.$$

因此

$$\alpha(D_m,\ D_j)\leqslant\varepsilon,\quad j,m>K.$$

由此可得 $\qquad \alpha(A_j,\ D_j)\leqslant\varepsilon,\quad j>K$，所以

$$\alpha(A,\ D_j)\leqslant\varepsilon,\quad j>K. \qquad (1.8\text{-}9)$$

用反证法可证存在$M\in\mathcal{N}$使$\alpha(A_m,\ A)\leqslant\varepsilon$，当$m>M$时. 倘若不然，将存在一序列$\{x_{n_i}\}$，对于所有的$i\in\mathcal{N}$，$x_{n_i}\in A_{n_i}$，且$d(x_{n_i},\ A)>\varepsilon$. 因为$A_1$是紧集，而且$x_{n_i}\in A_{n_i}\subset A_1$，所以序列$\{x_{n_i}\}$将有一个收敛的子列$\{y_i\}$. 设$y_i\to y\ (i\to\infty)$. 这将意味着$d(y,\ A)\geqslant\varepsilon$. 可证$y\in A$. 这二者是矛盾的. 因而存在$M\in\mathcal{N}$使$\alpha(A_m,\ A)\leqslant\varepsilon$，当$m>M$时.

因为$D_m\subset A_m$，所以当$m>M$时

$$\alpha(D_m,\ A)\leqslant\alpha(A_m,\ A)\leqslant\varepsilon. \qquad (1.8\text{-}10)$$

令$N=\max\{K,M\}$. 由式(1.8-9)与(1.8-10)得

$$h(D_i,\ A)\leqslant\varepsilon,\quad i>N.$$

这表明$\lim\limits_{i\to\infty}h(D_i,\ A)=0$，即$\lim\limits_{i\to\infty}D_i=A$.

由定理1.4.5与定理1.8.1可推得以下定理.

*定理 1.8.4 设$C\subset\mathscr{R}^n$为凸集，则下列条件等价：

1° C是凸体（即$\mathrm{int}(C)\neq\Phi$）；

2° C是凸代数体（即$C^i\neq\Phi$）；

3° $\dim(C)=n$.

证 要证明三个条件等价，只需证明"1°⇒2°"（1°是2°的充分条件），"2°⇒3°"，"3°⇒1°".

"1°⇒2°" 设C是凸体，则由定理1.4.5得$C^i=\mathrm{int}(C)\neq\Phi$，所以$C$是凸代数体.

"2°⇒3°" 设C是凸代数体，而$\dim(C)<n$. 即$\dim(\mathrm{aff}(C))<n$，因而$\mathrm{aff}(C)\neq\mathscr{R}^n$. 设$\boldsymbol{c}\in C^i$，$\boldsymbol{b}\in\mathscr{R}^n\setminus\mathrm{aff}(C)$，$\boldsymbol{m}$是通过$\boldsymbol{b}$，$\boldsymbol{c}$两点的直线. 显然$\boldsymbol{c}\in m\cap\mathrm{aff}(C)$. 若另有点$\boldsymbol{x}\in m\cap\mathrm{aff}(C)$，而$\boldsymbol{x}\neq\boldsymbol{c}$，则通过$\boldsymbol{c}$，$\boldsymbol{x}$的直线$m\subset\mathrm{aff}(C)$. 因而$\boldsymbol{b}\in\mathrm{aff}(C)$. 与假设矛盾. 所以$\dim(C)=n$.

"3°⇒1°" 设$\dim(C)=n$. 根据定理1.8.1 $C=\mathrm{co}(C)$至少包含一个n维单纯形$S=\mathrm{co}(\boldsymbol{a}^{(0)},\boldsymbol{a}^{(1)},\cdots,\boldsymbol{a}^{(n)})$. 可以证明

$$\boldsymbol{x}^{0}=\frac{1}{n+1}(\boldsymbol{a}^{(0)}+\boldsymbol{a}^{(1)}+\cdots+\boldsymbol{a}^{(n)})\in\mathrm{int}(S),$$

因而C是凸体.

由定理1.8.4可知，若$\dim(C)<n$，则$\mathrm{int}(C)=\Phi$，$C^i=\Phi$.

*§1.9 凸集的相对内部

有些凸集的内部是空的，然而把它作为其仿射包的子集考虑时，内部却是不空的. 例如，设$C=\mathrm{co}(\boldsymbol{e}^{(1)},\boldsymbol{e}^{(2)},\boldsymbol{e}^{(3)})$，其中$\boldsymbol{e}^{(1)}=(1,0,0)^T$，$\boldsymbol{e}^{(2)}=(0,1,0)^T$，$\boldsymbol{e}^{(3)}=(0,0,1)^T$. 在$\mathscr{R}^3$中，$\mathrm{int}(C)=\Phi$. 然而在仿射包$\mathrm{aff}(C)$中考虑时，$C$的内部是不空的. 因此引入凸集的相对内部的概念.

定义 1.9.1 设E为线性拓扑空间，$C\subset E$为凸集. 把C看作它

的仿射包 aff（C）的子集时C的内部称为**C 的相对内部**（relative interior），记为ri（C）.

定义 1.9.2 设E为线性拓扑空间，$C\subset E$为凸集. 把C看作它的仿射包aff（C）的子集时，C的边界称为**C的相对边界**，记为rb（C）.

对于前面提到的凸集$C=\mathrm{co}(e^{(1)}, e^{(2)}, e^{(3)})$，它的相对内部与相对边界分别是：

$$\mathrm{ri}(C)=\left\{x \middle| x=\sum_{i=1}^{3}\lambda_i e^{(i)},\ \lambda_i>0,\quad \sum_{i=1}^{3}\lambda_i=1\right\},$$

$$\mathrm{rb}(C)=[e^{(1)}, e^{(2)}]\cup[e^{(2)}, e^{(3)}]\cup[e^{(3)}, e^{(1)}].$$

关于凸集C的相对内部与相对边界有以下性质：

性质1 设$C\subset\mathscr{R}^n$. 若C为凸集，而且$\dim(C)=n$，则$\mathrm{ri}(C)=\mathrm{int}(C)$.

（因为$\mathrm{aff}(C)=\mathscr{R}^n$）

性质2 设$C\subset\mathscr{R}^n$. 若C为凸集，则

1° $\mathrm{ri}(C)$是凸集；

2° $\mathrm{rb}(C)=\overline{C}\setminus\mathrm{ri}(C)$.

性质3 设$C\subset\mathscr{R}^n$. 若C为有界凸集，则

$$\mathrm{ri}(C)\subset\mathrm{co}(\mathrm{rb}(C)).$$

证 若$C=\Phi$，或C为单点集$\{c\}$，命题显然成立.

设C中至少包含两个点. 再设$x\in\mathrm{ri}(C)$，$y\in C$，$y\neq x$，m为通过x，y的直线. 由于C为有界凸集，所以m与C的相对边界rb（C）必交于两个点p，q，使$x\in(p, q)$. 于是

$$x\in\mathrm{co}(rb(C)).$$

再由x的任意性得

$$\mathrm{ri}(C)\subset\mathrm{co}(\mathrm{rb}(C)).$$

性质4 设$C_1\subset C_2\subset E$，C_1，C_2均为凸集，而且$\mathrm{aff}(C_1)=\mathrm{aff}(C_2)$，则

$$\mathrm{ri}(C_1)\subset\mathrm{ri}(C_2).$$

性质4显然成立．但是，若去掉条件"$\mathrm{aff}(C_1)=\mathrm{aff}(C_2)$"，则命题不成立．例如，设$C_1=[e^{(1)}, e^{(2)}]$，$C_2=\mathrm{co}(e^{(1)}, e^{(2)}, e^{(3)})$，其中$e^{(1)}$，$e^{(2)}$，$e^{(3)}$为$\mathscr{R}^3$中的单位列向量（见p.44）．$C_1$，$C_2$均为凸集，$C_1\subset C_2$．然而$\mathrm{ri}(C_1)\cap\mathrm{ri}(C_2)=\Phi$.

根据凸集的相对内部的定义可得以下定理.

定理 1.9.1 设$C\subset\mathscr{R}^n$是凸集.

1° 若 $x\in\mathrm{ri}(C)$，$y\in\overline{C}$，则$(x, y)\subset\mathrm{ri}(C)$；

2° $\overline{\mathrm{ri}(C)}=\overline{C}$，$\mathrm{ri}(\overline{C})=\mathrm{ri}(C)$.

证 1° 把C看作它的仿射包$\mathrm{aff}(C)$的子集，然后应用定理1.4.1即得.

2° 把C看作它的仿射包$\mathrm{aff}(C)$的子集，然后应用定理1.4.5即得.

定理 1.9.2 设$C\subset\mathscr{R}^n$是凸集，$\mathrm{ri}(C)\neq\Phi$，则

1° $\mathrm{ri}(C)=\mathrm{ri}(\overline{C})=\mathrm{ri}(\mathrm{ri}(C))$；

2° $\mathrm{rb}(C)=\mathrm{rb}(\overline{C})=\mathrm{rb}(\mathrm{ri}(C))$；

3° $\mathrm{aff}(C)=\mathrm{aff}(\overline{C})=\mathrm{aff}(\mathrm{ri}(C))$；

4° $\dim(C)=\dim(\overline{C})=\dim(\mathrm{ri}(C))$.

（证明从略）

定理 1.9.3 设$C\subset\mathscr{R}^n$为非空凸集，$\dim(C)\geqslant 1$，则$\mathrm{ri}(C)\neq\Phi$.

（根据定义1.9.1与定理1.8.4可证）

定理 1.9.4 设$C\subset\mathscr{R}^n$为非空凸集．则点$x\in\mathrm{ri}(C)$的充分必要条件是：$\forall y\in C$存在$\mu>1$，使

$$(1-\mu)y+\mu x\in C \quad (1).$$

定理 1.9.5 设C，$D\subset\mathscr{R}^n$均为非空凸集，$\mathrm{ri}(C)\cap\mathrm{ri}(D)$

$\neq \Phi$. 则

1° $\overline{C \cap D} = \overline{C} \cap \overline{D}$;

2° $\mathrm{ri}(C \cap D) = \mathrm{ri}(C) \cap \mathrm{ri}(D)$.

证 1° 因为$C \cap D \subset \overline{C} \cap \overline{D}$，$\overline{C} \cap \overline{D}$为闭集，所以$\overline{C \cap D} \subset \overline{C} \cap \overline{D}$. 下面再证明反包含关系.

设$x \in \overline{C} \cap \overline{D}$，$b \in \mathrm{ri}(C) \cap \mathrm{ri}(D)$. 由定理1.9.1得$[b, x) \subset \mathrm{ri}(C) \cap \mathrm{ri}(D)$，因而$x \in \overline{\mathrm{ri}(C) \cap \mathrm{ri}(D)}$.
由x的任意性得

$$\overline{C} \cap \overline{D} \subset \overline{\mathrm{ri}(C) \cap \mathrm{ri}(D)} \subset \overline{C \cap D} \subset \overline{C} \cap \overline{D}, \quad (1.9\text{-}1)$$

所以

$$\overline{C \cap D} = \overline{C} \cap \overline{D}.$$

2° 由式(1.9-1)得 $\overline{\mathrm{ri}(C) \cap \mathrm{ri}(D)} = \overline{C \cap D}$，因而

$$\mathrm{ri}(\overline{C \cap D}) = \mathrm{ri}(\overline{\mathrm{ri}(C) \cap \mathrm{ri}(D)}).$$

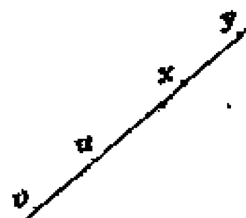


图1-6

依定理1.9.1得

$$\mathrm{ri}(\overline{C \cap D}) = \mathrm{ri}(C \cap D),$$
$$\mathrm{ri}(\overline{\mathrm{ri}(C) \cap \mathrm{ri}(D)}) = \mathrm{ri}(\mathrm{ri}(C) \cap \mathrm{ri}(D)),$$

所以

$$\mathrm{ri}(C \cap D) = \mathrm{ri}(\mathrm{ri}(C) \cap \mathrm{ri}(D)),$$

因而

$$\mathrm{ri}(C \cap D) \subset \mathrm{ri}(C) \cap \mathrm{ri}(D). \quad (1.9\text{-}2)$$

下面再证明反包含关系. 设$x \in \mathrm{ri}(C) \cap \mathrm{ri}(D)$. 由$\mathrm{ri}(C) \cap \mathrm{ri}(D) \neq \Phi$得$C \cap D \neq \Phi$,且$\dim(C \cap D) \geqslant 1$，根据定理1.9.3得$\mathrm{ri}(C \cap D) \neq \Phi$，所以可设$y \in \mathrm{ri}(C \cap D)$，因而$y \in C \cap D$. 若$x \neq y$，则在通过$x$，$y$的直线上必存在$u \in C$，$v \in D$ 使$x \in (y, u)$，$x \in (y, v)$（见图1-6）. 不妨设$(y, u) \subset (y, v)$,

因而$u\in C\cap D$. 依定理1.9.1得$[y, u)\subset \mathrm{ri}(C\cap D)$，因而$x\in \mathrm{ri}(C\cap D)$. 由$x$的任意性得

$$\mathrm{ri}(C)\cap \mathrm{ri}(D)\subset \mathrm{ri}(C\cap D). \qquad (1.9\text{-}3)$$

由式(1.9-2)与(1.9-3)得

$$\mathrm{ri}(C\cap D)=\mathrm{ri}(C)\cap \mathrm{ri}(D).$$

定理 1.9.6 设$C\subset\mathscr{R}^n$为非空凸集，T为由$\mathscr{R}^n$到$\mathscr{R}^m$的线性映射，则

1° $T(\overline{C})\subset\overline{TC}$;

2° $T(\mathrm{ri}(C))=\mathrm{ri}(TC)$.

证 1° 因为由$\mathscr{R}^n$到$\mathscr{R}^m$的线性映射T是连续的，而在连续线性映射之下闭包的象含于象的闭包中，即$T(\overline{C})\subset\overline{TC}$.

2° 因为C为凸集，所以$\mathrm{ri}(C)$也是凸集. 于是由1°得

$$T(\overline{\mathrm{ri}(C)})\subset\overline{T(\mathrm{ri}(C))}.$$

根据定理1.9.1 $\overline{\mathrm{ri}(C)}=\overline{C}$,因而$T(\overline{\mathrm{ri}(C)})=T(\overline{C})$. 显然$T(\overline{C})\supset TC\supset T(\mathrm{ri}(C))$. 于是

$$\overline{T(\mathrm{ri}(C))}\supset TC\supset T(\mathrm{ri}(C)),$$

因而 $\overline{TC}=\overline{T(\mathrm{ri}(C))}$. 因为$C$是凸集，$T$是线性映射，所以根据§1.1性质4 TC也是凸集. 再根据定理1.9.1得

$$\mathrm{ri}(TC)=\mathrm{ri}(\overline{TC})=\mathrm{ri}(\overline{T(\mathrm{ri}(C))})$$
$$=\mathrm{ri}(T(\mathrm{ri}(C))),$$

因此 $\mathrm{ri}(TC)\subset T(\mathrm{ri}(C))$. 下面再证反包含关系.

设$x\in T(\mathrm{ri}(C))$，即存在$x^{(1)}\in\mathrm{ri}(C)$，使$x=Tx^{(1)}$. 再设$y\in TC$，即$y^{(1)}\in C$，且$Ty^{(1)}=y$. 因为$x^{(1)}\in\mathrm{ri}(C)$，所以存在$\mu>1$，使

$$(1-\mu)y^{(1)}+\mu x^{(1)}\in C,$$

因而

$$T[(1-\mu)y^{(1)}+\mu x^{(1)}]\in TC,$$

即

$$(1-\mu)y+\mu x\in TC.$$

由定理1.9.4知，$x\in \mathrm{ri}(TC)$．由x的任意性得

$$T(\mathrm{ri}(C))\subset \mathrm{ri}(TC).$$

综上所述得 $T(\mathrm{ri}(C))=\mathrm{ri}(TC)$.

定理 1.9.7 设C，$D\subset\mathscr{R}^n$为凸集，且$\lambda\in\mathscr{R}$．则

1° $\mathrm{ri}(\lambda C)=\lambda\mathrm{ri}(C)$;

2° $\mathrm{ri}(C+D)=\mathrm{ri}(C)+\mathrm{ri}(D)$.

证 1° 设由$Tx=\lambda x$确定线性映射$T:\mathscr{R}^n\to\mathscr{R}^n$．再应用定理1.9.6即得$\mathrm{ri}(\lambda C)=\lambda\mathrm{ri}(C)$.

2° 设由$T(x, y)=x+y$确定线性映射$T:\mathscr{R}^n\times\mathscr{R}^n\to\mathscr{R}^n$．将定理1.9.6用于此线性映射得

$$\mathrm{ri}(C+D)=\mathrm{ri}(T(C\times D))=T(\mathrm{ri}(C\times D)),$$

可以证明 $\mathrm{ri}(C\times D)=\mathrm{ri}(C)\times\mathrm{ri}(D)$.于是

$$\begin{aligned}\mathrm{ri}(C+D)&=T(\mathrm{ri}(C\times D))\\&=T(\mathrm{ri}(C)\times\mathrm{ri}(D))\\&=\mathrm{ri}(C)+\mathrm{ri}(D),\end{aligned}$$

因而

$$\mathrm{ri}(C+D)=\mathrm{ri}(C)+\mathrm{ri}(D).$$

习　　题

1． 设A，$B\subset V$是凸集，$A\cup B$是否为凸集？举例说明之．

2． 设$C\subset V$为凸集，x^1，…，$x^k\in C$．试证明

$$\sum_{i=1}^{k}\lambda_i x^i\in C,\ \forall\lambda_i\geqslant 0,\ \sum_{i=1}^{k}\lambda_i=1.$$

3． 设E为赋范线性空间，$C\subset E$为非空闭凸集，$\alpha>0$．试证明集合

$$G=\{x\mid d(x, C)<\alpha\}$$

为非空开凸集．

*4. 试举出一个闭集$A\subset V$，而co(A)非闭集.

5. 设$A\subset V$为仿射集，$\boldsymbol{x}^{(1)}$, …, $\boldsymbol{x}^{(k)}\in A$. 试证明

$$\sum_{i=1}^{k}\lambda_i\boldsymbol{x}^{(i)}\in A,\ \forall\lambda_i\in\mathscr{R},\ \sum_{i=1}^{k}\lambda_i=1.$$

6. 设A为$m\times n$矩阵，$\boldsymbol{b}\in\mathscr{R}^m$，$\boldsymbol{b}\neq 0$. 试证明非齐次线性方程组的解集

$$M=\{\boldsymbol{x}\mid A\boldsymbol{x}=\boldsymbol{b}\}$$

是一个仿射集.

7. 试证明

$$\dim(\text{aff}(\boldsymbol{x}^{(1)},\ \cdots,\ \boldsymbol{x}^{(m)}))\leqslant m-1.$$

8. 设$A\subset\mathscr{R}^n$, $\dim(A)=k$. 若A中p个点$\boldsymbol{x}^{(1)}$, …, $\boldsymbol{x}^{(p)}$仿射无关，则$p\leqslant k+1$；若$p<k+1$，则A中存在点$\boldsymbol{x}^{(p+1)}$, …, $\boldsymbol{x}^{(k+1)}$与$\boldsymbol{x}^{(1)}$, …, $\boldsymbol{x}^{(p)}$仿射无关.

*9. 设$S=\text{co}(\boldsymbol{x}^{(0)},\ \boldsymbol{x}^{(1)},\ \cdots,\ \boldsymbol{x}^{(n)})$为一个$n$维单纯形，试证明

$$\text{int}(S)=\left\{\sum_{i=0}^{n}\lambda_i\boldsymbol{x}^{(i)}\ \middle|\ \lambda_i>0,\ \sum_{i=0}^{n}\lambda_i=1\right\}.$$

*10. 设$A\subset\mathscr{R}^n$是开集，$\boldsymbol{x}^{(0)}\in A$. 试证明存在$n$维单纯形$S$，使$S\subset A$，且$\boldsymbol{x}^{(0)}\in\text{int}(S)$.

11. 设$A=\{\boldsymbol{a}\}$，试证明$A^i=A=\{\boldsymbol{a}\}$.

12. 设Φ为空集. 下列二式是否都正确?

1° $\Phi+\{\boldsymbol{a}\}=\{\boldsymbol{a}\}$,　　2° $\Phi\cup\{\boldsymbol{a}\}=\{\boldsymbol{a}\}$.

*13. 设$C\subset V$为凸代数体，$\boldsymbol{x}^{(0)}\in C^i$，$0\overline{\in}C^i$. 试用关于$C$的Minkowski泛函$p(\boldsymbol{x})$来刻划$C^i$与$C^a$.

*14. 设C_1, $C_2\subset E$均为凸体. 试证明$\overline{C_1}=\overline{C_2}$的充分必要条件是

$$\text{int}(C_1)=\text{int}(C_2).$$

*15. 设C_1，…，$C_n \subset E$均为凸体，且

$$\bigcap_i \mathrm{int}(C_i) \neq \Phi,$$

试证明

$$\overline{\bigcap_i \mathrm{int}(C_i)} = \bigcap_i C_i .$$

16. 设K_1，K_2是包含点0的凸锥．试证明

$$K_1 + K_2 = \mathrm{co}(K_1 \cup K_2).$$

*17. 试证明具有Hausdorff距离的集族$\Omega(r)$（见定义1.8.2）是度量空间．

18. 设$K \subset E$是一个非空凸锥，试证明$\overline{K}$是包含点0的凸锥．

第二章　凸集分离定理

凸集分离定理是凸分析中最重要的成果之一．本章首先引入两个凸集被超平面H分离的概念，接着讨论在一般线性空间V以及线性拓扑空间E中两个凸集的分离定理，然后讨论有限维线性空间$\mathscr{R}^n$中两个凸集的分离定理，最后讨论凸集分离定理的应用．

§2.1　概念·引理

设C，$D\subset V$，$H=f^{-1}(\alpha)$是V中一张超平面，其中$f:V\to\mathscr{R}$为线性函数，并用$f(C)\geqslant\alpha$表示$f(\boldsymbol{x})\geqslant\alpha$，$\forall\,\boldsymbol{x}\in C$．

定义 2.1.1　若$f(C)\geqslant\alpha$，$f(D)\leqslant\alpha$，且C与D并不都含于H中，则称超平面H**分离**C，D（见图2-1（a））．

定义 2.1.2　若$f(C)>\alpha$，$f(D)<\alpha$，则称H严格分离C，D（见图2-1（b））．

定义 2.1.3　设C，$D\subset\mathscr{R}^n$为非空凸集，$B=\{\boldsymbol{x}\mid\|\boldsymbol{x}\|\leqslant1\}$．若存在$\varepsilon>0$，使$C+\varepsilon B$与$D+\varepsilon B$被超平面$\boldsymbol{p}^T\boldsymbol{x}=\alpha$严格分离，则称

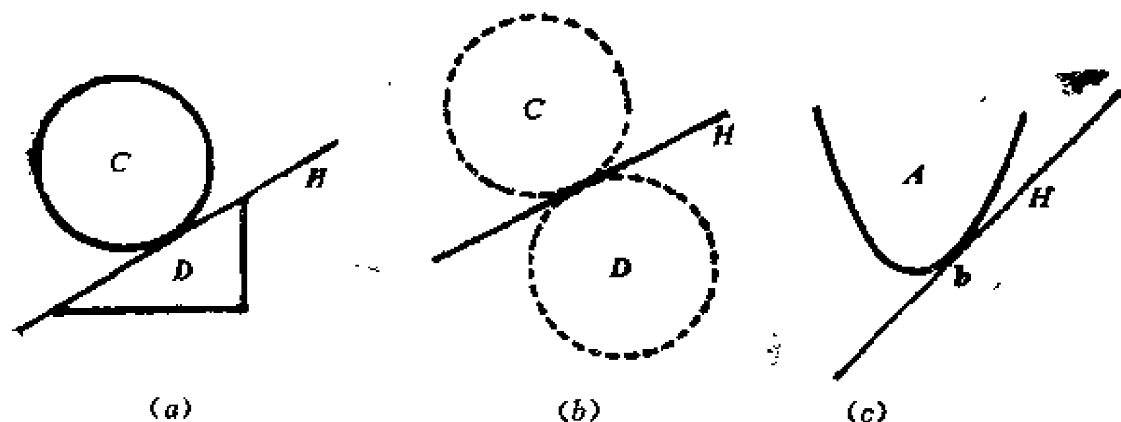


图 2-1

超平面$p^T x = \alpha$强分离C与D。

定义 2.1.4 设$A \subset V$, $b \in \mathrm{bd}(A)$, $H = f^{-1}(\alpha)$为V中一张超平面. 若$b \in H$, $f(A) \geqslant \alpha$(或$f(A) \leqslant \alpha$), 则称H为在点b对于集合A的一张**支撑超平面**(supporting hyperplane)(见图 2-1(c)).

若A的支撑超平面$H \supset A$, 则称H为平凡的支撑超平面; 反之, 称为非平凡的支持超平面. 重要的是非平凡的支撑超平面.

例1 设 $C = \{x \mid (x_1+1)^2 + (x_2-1)^2 \leqslant 2\}$,

$D = \{x \mid (x_1-1)^2 + (x_2+1)^2 \leqslant 2\}$,

则超平面H: $x_1 - x_2 = 0$ 分离C, D.

例2 设 $C = \{x \mid (x_1+1)^2 + (x_2-1)^2 < 2\}$,

$D = \{x \mid (x_1-1)^2 + (x_2+1)^2 < 2\}$,

则超平面H: $x_1 - x_2 = 0$ 严格分离C, D.

例3 设 $C = \{x \mid (x_1+1)^2 + (x_2-1)^2 < 1\}$,

$D = \{x \mid (x_1-1)^2 + (x_2+1)^2 < 1\}$,

则超平面H: $x_1 - x_2 = 0$ 强分离C, D.

例 4 设$A = \{x \mid x_2 \geqslant x_1^2 + 1\}$, $b = (1, 2)^T$, 则超平面H: $2x_1 - x_2 = 0$ 是在点b对A的一张非平凡的支撑超平面.

若超平面H强分离C, D, 则必严格分离C, D; 若H严格分离C, D, 则必分离C, D. 但是, 由H分离C, D不能推得H严格分离C, D(见例1). 由H严格分离C, D不能推得H强分离C, D(见例2).

引理 2.1.1 设A, $B \subset V$为凸集, $A \cap B = \Phi$.

1° 若 $x \overline{\in} A \cup B$, 则$\mathrm{co}(\{x\} \cup A) \cap B = \Phi$或$\mathrm{co}(\{x\} \cup B) \cap A = \Phi$;

2° 存在凸集C, $D \subset V$, 使$C \supset A$, $D \supset B$, $C \cap D = \Phi$, $C \cup D = V$.

***证** 1° 若 $x \overline{\in} A \cup B$, 而$\mathrm{co}(\{x\} \cup A) \cap B \neq \Phi$, $\mathrm{co}(\{x\} \cup B)$

$\cap A\neq\Phi$，则必存在$\boldsymbol{b}^{(1)}\in B$，$\boldsymbol{a}^{(1)}\in A$，$\lambda\in(0,1)$，使

$$\boldsymbol{b}^{(1)}=\lambda\boldsymbol{x}+(1-\lambda)\boldsymbol{a}^{(1)}.$$

与此同时还存在$\boldsymbol{a}^{(2)}\in A$，$\boldsymbol{b}^{(2)}\in B$，$\mu\in(0,1)$，使

$$\boldsymbol{a}^{(2)}=\mu\boldsymbol{x}+(1-\mu)\boldsymbol{b}^{(2)},$$

即$\boldsymbol{b}^{(1)}\in(\boldsymbol{x},\boldsymbol{a}^{(1)})$，$\boldsymbol{a}^{(2)}\in(\boldsymbol{x},\boldsymbol{b}^{(2)})$. 而$\boldsymbol{x}$，$\boldsymbol{a}^{(1)}$与$\boldsymbol{b}^{(2)}$构成一个三角形（见图2-2）. 线段$[\boldsymbol{a}^{(1)},\boldsymbol{a}^{(2)}]$与线段$[\boldsymbol{b}^{(1)},\boldsymbol{b}^{(2)}]$相交于一点$\boldsymbol{y}$，$\boldsymbol{y}\in A\cap B$. 这与假设矛盾. 于是命题得证.

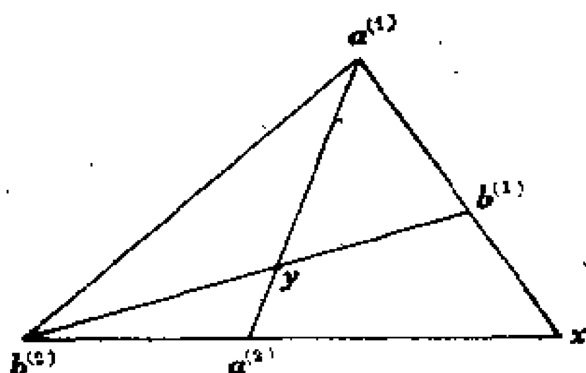


图 2-2

2° 设M是V的有序的凸子集对(P,Q)的全体，$P\supset A$，$Q\supset B$，$P\cap Q=\Phi$；在M中定义序关系"$\leqslant$"如下：

$(P_1,Q_1)\leqslant(P_2,Q_2)$，即$P_1\subset P_2$，$Q_1\subset Q_2$.

M的每个全序子集$\{(P_\alpha,Q_\alpha)\}$都有一个上界$(\bigcup_\alpha P_\alpha,\bigcup_\alpha Q_\alpha)$. 显然$\bigcup_\alpha P_\alpha\supset A$，$\bigcup_\alpha Q_\alpha\supset B$. 用反证法可得$(\bigcup_\alpha P_\alpha)\cap(\bigcup_\alpha Q_\alpha)=\Phi$. 因而$(\bigcup_\alpha P_\alpha,\bigcup_\alpha Q_\alpha)\in M$. 根据下面介绍的Zorn引理，$M$必有极大元$(C,D)$. $C\supset A$，$D\supset B$，$C\cap D=\Phi$. 若$C\cup D\neq V$，则由1°可知(C,D)不是极大元，因而必有$C\cup D=V$.

Zorn引理[10] 若序集M中每一全序子集都有上界，则M必有极大元.

引理2.1.2 设$B\subset\mathscr{R}^n$为凸集，并且是相对开的（即$\mathrm{ri}(B)=B$），$n\geqslant2$，$F\subset\mathscr{R}^n$为线性子空间，$0\leqslant\dim(F)\leqslant n-2$，且$F\cap B$

$=\Phi$，则$\mathscr{R}^n$中存在一张超平面H，$H\supset F$且$H\cap B=\Phi$.

（证明从略）

引理 2.1.3 设$B\subset\mathscr{R}^n$为凸集，并且是相对开的（即$\mathrm{ri}(B)=B$），$n\geqslant 2$，$0\overline{\in}B$，则$\mathscr{R}^n$中存在一张超平面H，包含点0，且$H\cap B=\Phi$.

证 令$F=\{0\}$，再由引理2.1.2即可得证.

§2.2 凸集分离定理

2.2.1 线性空间中的凸集分离定理

下面研究线性空间、线性拓扑空间及线性赋范空间中的分离定理（separation theorem）.

定理 2.2.1 设A，$B\subset V$为凸集，$A\neq\Phi$，$B^i\neq\Phi$，$A\cap B^i=\Phi$，则存在超平面H分离A，B.

证 根据引理2.1.1存在凸集C，$D\subset V$，$C\supset A$，$D\supset B^i$，$C\cap D=\Phi$，$C\cup D=V$. 若$\boldsymbol{x}\in B^i$，则$\boldsymbol{x}\in D$，因而$\boldsymbol{x}\overline{\in}C$. 而$C\subset C^a$，所以$B^i\cap C^a=\Phi$. 因为$B^i\neq\Phi$，所以存在$\boldsymbol{y}\in B^i$. 由以上分析可知，$\boldsymbol{y}\overline{\in}C^a$，因而$\boldsymbol{y}\overline{\in}C^a\cap D^a$. 根据定理1.5.2 $H=C^a\cap D^a$是一张超平面. 令$H=f^{-1}(\alpha)$，其中f为线性函数，$\alpha\in\mathscr{R}$. 因为$B^i\cap H=B^i\cap C^a\cap D^a=\Phi$，所以存在$\boldsymbol{b}\in B^i$，使$f(\boldsymbol{b})\neq\alpha$，不妨设$f(\boldsymbol{b})<\alpha$. 若存在$\boldsymbol{x}\in B$使$f(\boldsymbol{x})>\alpha$，则必存在$\boldsymbol{z}\in(\boldsymbol{x},\boldsymbol{b})$，使$f(\boldsymbol{z})=\alpha$（因为$f$是线性函数），因而$\boldsymbol{z}\in H$. 可是$[\boldsymbol{b},\boldsymbol{x})\subset B^i$（根据定理1.3.2），因而$\boldsymbol{z}\in B^i$. 这与$B^i\cap H=\Phi$矛盾. 所以$f(B)\leqslant\alpha$.

假设存在$\boldsymbol{x}\in A$，使$f(\boldsymbol{x})<\alpha$. 因为$A\subset C$，$\boldsymbol{b}\in B^i\subset D$，所以由定理1.3.3知存在$\boldsymbol{z}\in(\boldsymbol{b},\boldsymbol{x})\cap H$，即存在$\boldsymbol{z}\in(\boldsymbol{b},\boldsymbol{x})$，使$f(\boldsymbol{z})=\alpha$. 因为$f$是线性的，所以由$f(\boldsymbol{b})<\alpha$，$f(\boldsymbol{x})<\alpha$必将推得$f(\boldsymbol{z})<\alpha$，这与$f(\boldsymbol{z})=\alpha$矛盾，所以$f(A)\geqslant\alpha$.

因为存在$\boldsymbol{b}\in B$，使$f(\boldsymbol{b})<\alpha$，所以超平面H分离A，B.

注 若$f(b)>\alpha$，则可推得$f(B)\geqslant\alpha$，$f(A)\leqslant\alpha$.

定理 2.2.2 设A，$B\subset E$为凸集，$A\neq\Phi$，$\text{int}(B)\neq\Phi$，$A\cap\text{int}(B)=\Phi$，则E中存在一张闭超平面分离A，B.

证 因为$\text{int}(B)\neq\Phi$（即B为凸体），所以根据定理1.4.5得$B^i=\text{int}(B)$，因而$B^i\neq\Phi$. 由定理2.2.1知E中存在一张超平面$H=f^{-1}(\alpha)$分离A，B. 不妨设存在$b\in B$，使$f(b)>\alpha$，$f(B)\geqslant\alpha$，且$f(A)\leqslant\alpha$. 若存在$x^{(1)}\in\text{int}(B)$，使$f(x^{(1)})=\alpha$，则必存在$x^{(2)}\in B$，使$x^{(1)}\in(b, x^{(2)})$，且$f(x^{(2)})<\alpha$（因为$f(x)$是线性的）. 这与$f(B)\geqslant\alpha$矛盾，所以$f(\text{int}(B))>\alpha$. 令$D=\{x\mid f(x)>\alpha\}$，则$\text{int}(B)\subset D$. 由$\text{int}(B)\neq\Phi$及

$$\text{int}(B)=\text{int}(\text{int}(B))\subset\text{int}(D)$$

得$\text{int}(D)\neq\Phi$. 由定理1.5.4知超平面H是闭的.

定理 2.2.3 设$B\subset E$为凸体，$M\subset E$为仿射集，$M\neq\Phi$，且$M\cap\text{int}(B)=\Phi$，则E中存在一张包含M的闭超平面H分离B与M.

证 因为M也是凸集，所以由定理2.2.2知，E中存在一张闭超平面$H_1=f^{-1}(\alpha)$，分离B与M. 可设$f(B)\geqslant\alpha$，$f(M)\leqslant\alpha$，且存在$b\in B$，使$f(b)>\alpha$. 若$M\not\subset H_1$，则存在$m\in M$，使$f(m)=\beta<\alpha$. 若存在$t\in M$，$f(t)\neq\beta$，则必存在$\lambda\in\mathscr{R}$使

$$\begin{aligned}f(\lambda m+(1-\lambda)t)&=f(t)+\lambda[f(m)-f(t)]\\&=f(t)+\lambda[\beta-f(t)]>\alpha,\end{aligned}$$

这与 $\lambda m+(1-\lambda)t\in M$ 矛盾，所以$f(M)=\beta$，即超平面 $H=f^{-1}(\beta)$包含M. 而$f(B)\geqslant\alpha>\beta$,所以超平面$H$分离$B$与$M$. 因为$H_1=f^{-1}(\alpha)$是闭超平面，所以$f(x)$是连续函数，因而$H=f^{-1}(\beta)$也是闭超平面（根据定理1.5.4）.

推论 设$C\subset E$为一个凸体，$b\in\text{bd}(C)$. 则存在一张在点b对C的非平凡的闭的支撑超平面.

定理 2.2.4 设E是赋范线性空间，$C\subset E$为非空闭凸集，a

$\overline{\in} C$．则E中存在一张严格分离$\boldsymbol{a}$与C的闭超平面H．

***证**　因为C是闭集，$\boldsymbol{a}\overline{\in} C$，所以

$$\sigma=d(\boldsymbol{a},\ C)=\inf_{x\in C}\ \|\boldsymbol{x}-\boldsymbol{a}\|>0\ .$$

易知　$B=\{\boldsymbol{x}\mid \|\boldsymbol{x}-\boldsymbol{a}\|<\frac{1}{2}\sigma\}$　为非空开凸集．另外

$$G=\left\{\boldsymbol{x}\mid d(\boldsymbol{x},\ C)<\frac{1}{2}\sigma\right\} \qquad (2.2\text{-}1)$$

也是非空开凸集（见第一章习题3），而且$B\cap G=\Phi$．倘若不然，存在$\boldsymbol{x}^{(1)}\in B\cap G$．于是

$$\|\boldsymbol{x}^{(1)}-\boldsymbol{a}\|<\frac{1}{2}\sigma,\ d(\boldsymbol{x}^{(1)},\ C)<\frac{1}{2}\sigma.$$

而

$$d(\boldsymbol{a},\ C)\leqslant\|\boldsymbol{x}^{(1)}-\boldsymbol{a}\|+d(\boldsymbol{x}^{(1)},\ C), \qquad (2.2\text{-}2)$$

于是得

$$\sigma<\frac{1}{2}\sigma+\frac{1}{2}\sigma=\sigma.$$

这是不可能的．所以$B\cap G=\Phi$．因为B为非空开凸集，所以$B=\mathrm{int}(B)\neq\Phi$.根据定理2.2.2 E中存在一张闭超平面$H=f^{-1}(\alpha)$分离B与G．可设$f(B)\geqslant\alpha$，$f(G)\leqslant\alpha$．因为B，G均为开集，所以$f(B)>\alpha$，$f(G)<\alpha$．而$\boldsymbol{a}\in B$，$C\subset G$，所以超平面H严格分离$\boldsymbol{a}$与C．

定理 2.2.5　设E是赋范线性空间,$K\subset E$为非空闭凸锥，$\boldsymbol{b}\overline{\in} K$，则存在分离$\boldsymbol{b}$与$K$且包含点$\boldsymbol{0}$的闭超平面$H$．

（证明留给读者）

定理 2.2.6　设C，$D\subset E$均为非空开凸集，$C\cap D=\Phi$，则E中存在一张严格分离C，D的闭超平面．

***证**　因为$C\neq\Phi$，$\mathrm{int}(D)=D$，$C\cap\mathrm{int}(D)=C\cap D=\Phi$，根

据定理2.2.2 E 中存在闭超平面 $H=f^{-1}(\alpha)$ 分离 C 与 D. 不妨设 $f(C)\geqslant\alpha$, $f(D)\leqslant\alpha$. 因为 C 是开集，所以 H 不能包含 C，因面存在点 $\boldsymbol{c}\in C$, 使 $f(\boldsymbol{c})>\alpha$. 若存在 $\boldsymbol{x}^{(0)}\in C$, 使 $f(\boldsymbol{x}^{0})=\alpha$, 则必存在 $\boldsymbol{x}^{(1)}\in C$, $\boldsymbol{x}^{(0)}\in(\boldsymbol{c}, \boldsymbol{x}^{(1)})$, $f(\boldsymbol{x}^{(1)})<\alpha$. 这与 $f(C)\geqslant\alpha$ 矛盾. 于是得 $f(C)>\alpha$. 同理可得 $f(D)<\alpha$. 即闭超平面严格分离 C, D.

2.2.2 $\mathscr{R}^n$ 中的凸集分离定理

面有限维线性空间应用最广泛,故仅对有限维线性空间 $\mathscr{R}^n$ 中的凸集分离定理作进一步的研究.

定理 2.2.7 (Kirchberger定理) 设 A, B 为 $\mathscr{R}^n$ 中的有限集. 假设对于每个由至多 $n+2$ 个点组成的 $A\cup B$ 的子集 C 都存在一张超平面，严格分离 $C\cap A$ 与 $C\cap B$，则存在一张超平面严格分离 A 与 B.

证 若 $A\cup B$ 中包含的点数 $s\leqslant n+2$，命题显然成立，所以假设 $s>n+2$. 引入有限个集合：

$A(\boldsymbol{x})=\{(\boldsymbol{y}, \beta)\,|\,(\boldsymbol{y}, \beta)\in\mathscr{R}^n\times\mathscr{R}, \ \boldsymbol{y}^T\boldsymbol{x}>\beta\}$, $\forall\,\boldsymbol{x}\in A$;

$B(\boldsymbol{x})=\{(\boldsymbol{y}, \beta)\,|\,(\boldsymbol{y}, \beta)\in\mathscr{R}^n\times\mathscr{R}, \ \boldsymbol{y}^T\boldsymbol{x}<\beta\}$, $\forall\,\boldsymbol{x}\in B$.

容易验证 $A(\boldsymbol{x})$, $B(\boldsymbol{x})$ 都是 $\mathscr{R}^{n+1}=\mathscr{R}^n\times\mathscr{R}$ 中的非空凸集. 设 P 为由所有的 $A(\boldsymbol{x})$ 与 $B(\boldsymbol{x})$ 组成的集族. 对于 P 中任意 $n+2$ 个凸集 $A(\boldsymbol{x}^{(1)})$, $\cdots$, $A(\boldsymbol{x}^{(t)})$, $B(\boldsymbol{x}^{(t+1)})$, $\cdots$, $B(\boldsymbol{x}^{(n+2)})$, 令

$$C=\{\boldsymbol{x}^{(1)}, \cdots, \boldsymbol{x}^{(t)}, \boldsymbol{x}^{(t+1)}, \cdots, \boldsymbol{x}^{(n+2)}\},$$

于是

$$C\cap A=\{\boldsymbol{x}^{(1)}, \cdots, \boldsymbol{x}^{(t)}\}, \quad C\cap B=\{\boldsymbol{x}^{(t+1)}, \cdots, \boldsymbol{x}^{(n+2)}\}.$$

根据假设，存在超平面严格分离 $C\cap A$ 与 $C\cap B$，即存在 $\boldsymbol{b}\in\mathscr{R}^n$, $\boldsymbol{b}\neq 0$, $a\in\mathscr{R}$, 使

$$\boldsymbol{b}^T\boldsymbol{x}^{(i)}>a, \quad i=1, \cdots, t;$$

$$\boldsymbol{b}^T\boldsymbol{x}^{(j)}<a, \quad j=t+1, \cdots, n+2.$$

因为 $\boldsymbol{x}^i\in A$, $\boldsymbol{x}^j\in B$, 所以 $\mathscr{R}^n\times\mathscr{R}$ 中的点 $(\boldsymbol{b}, a)$ 属于 $A(\boldsymbol{x}^{(1)})$, $\cdots$, $A(\boldsymbol{x}^{(t)})$, $B(\boldsymbol{x}^{(t+1)})$, $\cdots$, $B(\boldsymbol{x}^{(n+2)})$, 因而属于它们的交集.

根据Helly定理（定理1.8 2）P的所有集合有一个非空的交，即存在公共点$(\boldsymbol{a},\ \gamma)\in\mathscr{R}^n\times\mathscr{R}$，使$\boldsymbol{a}^T\boldsymbol{x}>\gamma$，$\forall\boldsymbol{x}\in A$；$\boldsymbol{a}^T\boldsymbol{x}<\gamma$，$\forall\boldsymbol{x}\in B$，因而超平面$H=\{\boldsymbol{x}\mid\boldsymbol{a}^T\boldsymbol{x}=\gamma\}$严格分离$A$与$B$.

定理 2.2.8 设C_1，$C_2\subset\mathscr{R}^n$为非空凸集. 存在超平面分离C_1与C_2的充分必要条件是：存在向量$\boldsymbol{b}\in\mathscr{R}^n$满足

$$1^\circ\quad \inf\{\boldsymbol{b}^T\boldsymbol{x}\mid\boldsymbol{x}\in C_1\}\geqslant\sup\{\boldsymbol{b}^T\boldsymbol{x}\mid\boldsymbol{x}\in C_2\},\tag{2.2-3}$$

$$2^\circ\quad \sup\{\boldsymbol{b}^T\boldsymbol{x}\mid\boldsymbol{x}\in C_1\}>\inf\{\boldsymbol{b}^T\boldsymbol{x}\mid\boldsymbol{x}\in C_2\}.\tag{2.2-4}$$

证 充分性. 设向量$\boldsymbol{b}\in\mathscr{R}^n$满足条件$1^\circ$与$2^\circ$，因而$\boldsymbol{b}\neq\boldsymbol{0}$. 选$\beta$满足

$$\inf\{\boldsymbol{b}^T\boldsymbol{x}\mid\boldsymbol{x}\in C_1\}\geqslant\beta\geqslant\sup\{\boldsymbol{b}^T\boldsymbol{x}\mid\boldsymbol{x}\in C_2\},$$

于是得$\boldsymbol{b}^T\boldsymbol{x}\geqslant\beta$，$\forall\boldsymbol{x}\in C_1$；$\boldsymbol{b}^T\boldsymbol{x}\leqslant\beta$，$\forall\boldsymbol{x}\in C_2$. 由$2^\circ$可知，$C_1$与$C_2$不能都包含在超平面$H$：$\boldsymbol{b}^T\boldsymbol{x}=\beta$中，所以超平面$H$分离$C_1$，$C_2$.

必要性. 设超平面H：$\boldsymbol{b}^T\boldsymbol{x}=\alpha$分离$C_1$与$C_2$. 不妨设$\boldsymbol{b}^T\boldsymbol{x}\geqslant\alpha$，$\forall\boldsymbol{x}\in C_1$；$\boldsymbol{b}^T\boldsymbol{x}\leqslant\alpha$，$\forall\boldsymbol{x}\in C_2$. 因而条件$1^\circ$成立. 因为$H$分离$C_1$与$C_2$，所以至少存在一点$\boldsymbol{x}\in C_1$，使$\boldsymbol{b}^T\boldsymbol{x}>\alpha$，或至少存在一点$\boldsymbol{x}\in C_2$，使$\boldsymbol{b}^T\boldsymbol{x}<\alpha$. 于是条$2^\circ$成立.

***定理 2.2.9** 设C，$D\subset\mathscr{R}^n$为非空凸集，则$\mathscr{R}^n$中存在一张超平面H分离C与D的充分必要条件是

$$\mathrm{ri}(C)\cap\mathrm{ri}(D)=\Phi.\tag{2.2-5}$$

证 当$n=1$时命题显然正确. 设$n\geqslant2$，$A=C-D$. 因为C，D为非空凸集，所以A为非空凸集. 由定理1.9.7得$\mathrm{ri}(A)=\mathrm{ri}(C)-\mathrm{ri}(D)$. 因此式（2.2-5）等价于$\boldsymbol{0}\overline{\in}\mathrm{ri}(A)$. 充分性. 设条件（2.2-5）成立. 令$B=\mathrm{ri}(A)$，$B$是相对开的. 因为$A$是非空凸集，所以$B$也是非空凸集，而且$\boldsymbol{0}\overline{\in}B$. 由引理2.1.3知，$\mathscr{R}^n$中存在一张包含点$\boldsymbol{0}$的超平面$H$：$\boldsymbol{p}^T\boldsymbol{x}=0$，其中$\boldsymbol{p}\in\mathscr{R}^n$，$\boldsymbol{p}\neq\boldsymbol{0}$，$H\cap B=\Phi$. 不妨设$\boldsymbol{p}^T\boldsymbol{x}>0$，$\forall\boldsymbol{x}\in B$（若存在$\boldsymbol{x}^{(1)}$，$\boldsymbol{x}^{(2)}\in B$，使$\boldsymbol{p}^T\boldsymbol{x}^{(1)}>0$，$\boldsymbol{p}^T\boldsymbol{x}^{(2)}<0$，则必存在点$\boldsymbol{x}\in B$，使$\boldsymbol{p}^T\boldsymbol{x}=0$. 这与$H\cap B=\Phi$矛盾）. 由此得$\boldsymbol{p}^T\boldsymbol{x}\geqslant0$，$\forall\boldsymbol{x}\in A$. 因为$A=C-D$，所以

$$\boldsymbol{p}^T\boldsymbol{z}\geqslant\boldsymbol{p}^T\boldsymbol{y},\ \forall\boldsymbol{z}\in C,\ \forall\boldsymbol{y}\in D.$$

令$\beta=\inf\{\boldsymbol{p}^T\boldsymbol{z}\mid\boldsymbol{z}\in C\}$，则$\boldsymbol{p}^T\boldsymbol{z}\geqslant\beta$，$\forall\boldsymbol{z}\in C$；$\boldsymbol{p}^T\boldsymbol{y}\leqslant\beta$，$\forall\boldsymbol{y}\in D$
再由$\boldsymbol{p}^T\boldsymbol{x}>0$，$\forall\boldsymbol{x}\in B$知：存在$\boldsymbol{c}\in C$，$\boldsymbol{d}\in D$，使$\boldsymbol{p}^T(\boldsymbol{c}-\boldsymbol{d})>0$．因而必有$\boldsymbol{p}^T\boldsymbol{c}>\beta$或$\boldsymbol{p}^T\boldsymbol{d}<\beta$，即$C$，$D$不能都含于$H$中．于是超平面$H$：$\boldsymbol{p}^T\boldsymbol{x}=\beta$分离$C$与$D$.

必要性．设H：$\boldsymbol{p}^T\boldsymbol{x}=\alpha$分离$C$与$D$．不妨设$\boldsymbol{p}^T\boldsymbol{x}\geqslant\alpha$，$\forall\boldsymbol{x}\in C$；$\boldsymbol{p}^T\boldsymbol{y}\leqslant\alpha$，$\forall\boldsymbol{y}\in D$，并且存在$\boldsymbol{c}\in C$，使$\boldsymbol{p}^T\boldsymbol{c}>\alpha$．因而$\boldsymbol{p}^T\boldsymbol{x}\geqslant 0$，$\forall\boldsymbol{x}\in A$，并且存在$\boldsymbol{a}\in A$，使$\boldsymbol{p}^T\boldsymbol{a}>0$．若点$0\in ri(A)$，则存在$\delta>0$，使$[-\delta\boldsymbol{a},\ \delta\boldsymbol{a}]\subset A$，因而

$$\boldsymbol{p}^T(-\delta\boldsymbol{a})=-\delta\boldsymbol{p}^T\boldsymbol{a}\geqslant 0\ .$$

这与$\boldsymbol{p}^T\boldsymbol{a}>0$矛盾．以所点$0\overline{\in}\mathrm{ri}(A)$，即式（2.2−5）成立.

***定理 2.2.10** 设$C\subset\mathscr{R}^n$为凸集，$\mathrm{ri}(C)\neq\Phi$，$\boldsymbol{b}\in\mathrm{rb}(C)$，则$\mathscr{R}^n$中存在一张在点$\boldsymbol{b}$对$C$的非平凡的闭的支撑超平面.

证 在$\mathrm{aff}(C)$中应用定理2.2.3的推论即得.

定理 2.2.11 设$C_1,C_2\subset\mathscr{R}^n$为非空凸集．存在超平面$H$强分离$C_1$，$C_2$的充分必要条件是：存在非零向量$\boldsymbol{p}\in\mathscr{R}^n$，使

$$\inf\{\boldsymbol{p}^T\boldsymbol{x}\mid\boldsymbol{x}\in C_1\}>\sup\{\boldsymbol{p}^T\boldsymbol{x}\mid\boldsymbol{x}\in C_2\}.\qquad(2.2\text{-}6)$$

证 必要性．设超平面H：$\boldsymbol{p}^T\boldsymbol{x}=\alpha$强分离$C_1$，$C_2$，即存在$\varepsilon>0$，使$C_1+\varepsilon B$与$C_2+\varepsilon B$被$H$严格分离，其中$B$为单位球，亦即

$$\boldsymbol{p}^T(\boldsymbol{x}+\varepsilon\boldsymbol{y})>\alpha,\ \forall\boldsymbol{x}\in C_1,\ \forall\boldsymbol{y}\in B;\qquad(2.2\text{-}7)$$

$$\boldsymbol{p}^T(\boldsymbol{x}+\varepsilon\boldsymbol{y})<\alpha,\ \forall\boldsymbol{x}\in C_2,\ \forall\boldsymbol{y}\in B.\qquad(2.2\text{-}8)$$

由式（2.2−7）得

$$\inf\{\boldsymbol{p}^T\boldsymbol{x}\mid\boldsymbol{x}\in C_1\}>\inf\{\boldsymbol{p}^T(\boldsymbol{x}+\varepsilon\boldsymbol{y})\mid\boldsymbol{x}\in C_1,\ \boldsymbol{y}\in B\}\geqslant\alpha.$$

同理得

$\sup\{\boldsymbol{p}^T\boldsymbol{x}\mid\boldsymbol{x}\in C_2\}<\alpha$．所以式（2.2−6）成立.

充分性．设存在非零向量$\boldsymbol{p}\in\mathscr{R}^n$，使式（2.2−6）成立．令$m_1=\inf\{\boldsymbol{p}^T\boldsymbol{x}\mid\boldsymbol{x}\in C_1\}$，$m_2=\sup\{\boldsymbol{p}^T\boldsymbol{x}\mid\boldsymbol{x}\in C_2\}$，$\alpha=\frac{1}{2}(m_1+$

m_2)，$\beta=\frac{1}{2}(m_1-m_2)$，则

$$p^T x \geqslant \alpha+\beta,\ \forall x \in C_1,\ p^T x \leqslant \alpha-\beta,\ \forall x \in C_2.$$

因为单位球B有界，所以$p^T y$，$y \in B$，有界．取充分小的正数ε，可使

$$p^T(x+\varepsilon y)=p^T x+\varepsilon p^T y>\alpha,\ \forall x \in C_1,\ \forall y \in B;$$

$$p^T(x+\varepsilon y)=p^T x+\varepsilon p^T y<\alpha,\ \forall x \in C_2,\ \forall y \in B.$$

这表明超平面H：$p^T x=\alpha$严格分离$C_1+\varepsilon B$与$C_2+\varepsilon B$，因而超平面H强分离C_1，C_2．

§2.3 分离定理的应用

凸分析在理论与实践上都有广泛的应用．限于篇幅，本节只讨论凸集分离定理在理论上的某些应用．

*2.3.1 紧凸集的表示

凸多面体中的每个点都可用其顶点（即极点）的凸组合表示．那么一般紧凸集中的点能否用它的极点的凸组合表示？关于这个问题有如下定理．

定理 2.3.1 设$C \subset \mathscr{R}^n$为非空紧凸集．则

1° C的极点集$J \neq \Phi$；

2° $C=\mathrm{co}(J)$；

3° 若$\dim(C)=k$，则$\forall x \in C$，存在$m(\leqslant k+1)$个极点，使x为这m个极点的凸组合．

证 若$\mathrm{din}(C)=0$，则C为单点集$\{c\}$．命题显然成立．所以设$\dim(C) \geqslant 1$．根据定理1.9.3 $\mathrm{ri}(C) \neq \Phi$．假设对于维数小于k的非空紧凸集命题正确，考虑维数等于k的非空紧凸集C．因为C非空有界，所以$\mathrm{rb}(C) \neq \Phi$，可以设$b \in \mathrm{rb}(C)$．根据定理2.2.10可设H：$p^T x=\beta$为在点b对于C的一张非平凡的支撑超平面，$p^T x \geqslant \beta$，$\forall x \in C$．因为C为紧凸集，而且$b \in H \cap C$，所以$H \cap C$也是非空

紧凸集．因为$C \not\subseteq H$，所以$\dim(H \cap C) < k$．根据假设，对于$H \cap C$命题正确．它的极点集$J_1 \neq \Phi$，而且$H \cap C = \mathrm{co}(J_1)$．

设点$\boldsymbol{a} \in J_1$，因而$\boldsymbol{a} \in H \cap C$．于是$\boldsymbol{p}^T\boldsymbol{a} = \beta$．若存在$\boldsymbol{x}, \boldsymbol{y} \in C$，$\lambda \in (0, 1)$，使$\boldsymbol{a} = \lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}$，则

$$\boldsymbol{p}^T\boldsymbol{a} = \lambda\boldsymbol{p}^T\boldsymbol{x} + (1-\lambda)\boldsymbol{p}^T\boldsymbol{y} = \beta.$$

因为$\boldsymbol{p}^T\boldsymbol{x} \geqslant \beta$，$\boldsymbol{p}^T\boldsymbol{y} \geqslant \beta$，所以$\boldsymbol{p}^T\boldsymbol{x} = \boldsymbol{p}^T\boldsymbol{y} = \beta$．因而$\boldsymbol{x}$，$\boldsymbol{y} \in H$，于是$\boldsymbol{x}$，$\boldsymbol{y} \in H \cap C$．因为$\boldsymbol{a} \in J_1$，所以$\boldsymbol{x} = \boldsymbol{y} = \boldsymbol{a}$．因而$\boldsymbol{a}$为$C$的极点．于是$C$的极点集$J \neq \Phi$，而且$J_1 \subset J$．

因为C是闭的，$\boldsymbol{b}$为$\mathrm{rb}(C)$上任一点，而且$\boldsymbol{b} \in H \cap C$，所以$\boldsymbol{b} \in \mathrm{co}(J_1) \subset \mathrm{co}(J)$．由此得$\mathrm{rb}(C) \subset \mathrm{co}(J)$．

由§1.9性质3得$\mathrm{ri}(C) \subset \mathrm{co}(\mathrm{rb}(C))$，而$\mathrm{co}(\mathrm{rb}(C)) \subset \mathrm{co}(J)$，所以$\mathrm{ri}(C) \subset \mathrm{co}(J)$．前面已得$\mathrm{rb}(C) \subset \mathrm{co}(J)$，于是$C \subset \mathrm{co}(J)$．而$\mathrm{co}(J) \subset C$，所以$C = \mathrm{co}(J)$．

由以上推导可知，当$\dim(C) = 0, 1, \cdots, n$时，1°与2°都成立．根据2°与定理1.8.1可证3°（把J看作定理1.8.1中的集合A）．

2.3.2 凸多胞形的表示

下面讨论凸多胞形（不一定是紧集）的表示问题．

定理 2.3.2 设有$\mathscr{R}^n$中的凸多胞形$S = \{\boldsymbol{x} \mid A\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geqslant 0\} \neq \Phi$，其中$A$为$m \times n$矩阵，秩为$m$，$\boldsymbol{b} \in \mathscr{R}^m$．再设$\boldsymbol{x}^{(1)}$，…，$\boldsymbol{x}^{(k)}$是$S$的极点，$\boldsymbol{d}^1$，…，$\boldsymbol{d}^l$为$S$的极方向，集合$\Lambda$如下：

$$\Lambda = \left\{ \sum_{i=1}^{k} \lambda_i \boldsymbol{x}^{(i)} + \sum_{j=1}^{l} \mu_j \boldsymbol{d}^{(j)} \,\middle|\, \sum_{i=1}^{k} \lambda_i = 1, \ \lambda_i, \ \mu_j \geqslant 0 \right\} \tag{2.3-1}$$

则$S = \Lambda$．

证 因为S是凸集，$\boldsymbol{x}^{(i)} \in S$，所以$\sum_{i=1}^{k} \lambda_i \boldsymbol{x}^{(i)} \in S$，其中$\lambda_i \geqslant 0$，$\sum_{i=1}^{k} \lambda_i = 1$．又因为$\boldsymbol{d}^{(j)}$是$S$的极方向，所以$\Lambda \subset S$．以下再证

$S \subset \Lambda$.

因为$S \neq \Phi$，所以S至少有一个极点，因而$\Lambda \neq \Phi$. 可以验证Λ是闭凸集. 假设存在$\boldsymbol{z} \in S$，$\boldsymbol{z} \overline{\in} \Lambda$，则根据定理2.2.4存在一张闭超平面严格分离$\boldsymbol{z}$与$\Lambda$. 因而可以设存在非零向量$\boldsymbol{p} \in \mathscr{R}^n$，$\alpha \in \mathscr{R}$，使

$$\boldsymbol{p}^T \boldsymbol{z} > \alpha, \tag{2.3-2}$$

$$\boldsymbol{p}^T \left(\sum_{i=1}^{k} \lambda_i \boldsymbol{x}^{(i)} + \sum_{j=1}^{l} \mu_j \boldsymbol{d}^{(j)} \right) < \alpha, \tag{2.3-3}$$

其中$\lambda_i \geqslant 0$，$\sum_{i=1}^{k} \lambda_i = 1$，$\mu_j \geqslant 0$，$j = 1, \cdots, l$.

因为当$\mu_j \geqslant 0$时式（2.3-3）恒成立，所以$\boldsymbol{p}^T \boldsymbol{d}^{(j)}$不能为正，即

$$\boldsymbol{p}^T \boldsymbol{d}^{(j)} \leqslant 0, \quad j = 1, \cdots, l. \tag{2.3-4}$$

在式（2.3-3）中适当选择系数λ_i与μ_j，可得

$$\boldsymbol{p}^T \boldsymbol{x}^{(i)} < \alpha, \quad i = 1, \cdots, k. \tag{2.3-5}$$

于是由式（2.3-2）与式（2.3-5）得

$$\boldsymbol{p}^T \boldsymbol{z} > \max_{1 \leqslant i \leqslant k} \{ \boldsymbol{p}^T \boldsymbol{x}^{(i)} \} = \boldsymbol{p}^T \hat{\boldsymbol{x}}, \tag{2.3-6}$$

其中$\hat{\boldsymbol{x}} \in \{ \boldsymbol{x}^{(1)}, \cdots, \boldsymbol{x}^{(k)} \}$，因而

$$\boldsymbol{p}^T \boldsymbol{z} - \boldsymbol{p}^T \hat{\boldsymbol{x}} > 0 \tag{2.3-7}$$

因为$\hat{\boldsymbol{x}}$为极点，所以由定理1.7.1知$A = [B, N]$，其中B为$m \times m$的非奇异矩阵，而且

$$\hat{\boldsymbol{x}} = \begin{pmatrix} \hat{\boldsymbol{x}}_B \\ \hat{\boldsymbol{x}}_N \end{pmatrix} = \begin{pmatrix} B^{-1} \boldsymbol{b} \\ 0 \end{pmatrix} \geqslant 0 \tag{2.3-8}$$

将$\boldsymbol{p}$与$\boldsymbol{z}$作相应的划分，即

$$\boldsymbol{p} = \begin{pmatrix} \boldsymbol{p}_B \\ \boldsymbol{p}_N \end{pmatrix}, \quad \boldsymbol{z} = \begin{pmatrix} \boldsymbol{z}_B \\ \boldsymbol{z}_N \end{pmatrix}. \tag{2.3-9}$$

因为$z\in S$，所以$Az=Bz_B+Nz_N=b$，$z\geqslant 0$．由此可得

$$z_B=B^{-1}b-B^{-1}Nz_N. \tag{2.3-10}$$

将式（2.3-8）～（2.3-10）代入式（2.3-7）得

$$\begin{aligned}p^Tz-p^T\hat{x}&=p_B^T(B^{-1}b-B^{-1}Nz_N)+p_N^Tz_N-p_B^TB^{-1}b\\&=(p_N^T-p_B^TB^{-1}N)z_N>0\ .\end{aligned}$$

因为$z_N\geqslant 0$，所以至少有一个$j\geqslant m+1$，使$z_j>0$而且$p_j-p_B{}^TB^{-1}a^{(j)}>0$，其中$a^{(j)}$为$A$的第$j$列．用反证法可证$y^{(j)}=B^{-1}a^{(j)}$必有正分量．设$b'=B^{-1}b$，

$$\lambda=\min_{1\leqslant i\leqslant m}\left\{\frac{b_i'}{y_i^{(j)}}\middle|y_i^{(j)}>0\right\}=\frac{b_r'}{y_r^{(j)}}. \tag{2.3-11}$$

假设$B^{-1}b>0$，则$\lambda>0$．再设

$$x=\begin{pmatrix}b'\\0\end{pmatrix}+\lambda\begin{pmatrix}-y^{(j)}\\e^{(j)}\end{pmatrix}, \tag{2.3-12}$$

其中$e^{(j)}=\mathscr{R}^{n-m}$，相应于$a^{(j)}$的分量为1，其余分量为零．

x至多有m个正分量，第j个分量为λ，而第r个分量为零，$x\geqslant 0$，而且

$$\begin{aligned}Ax&=B(b'-\lambda y^{(j)})+\lambda Ne^{(j)}\\&=B(B^{-1}b-\lambda B^{-1}a^{(j)})+\lambda a^{(j)}=b,\end{aligned}$$

所以$x\in S$．还可以证明x是S的极点，因而$x\in\{x^{(1)},\ \cdots,\ x^{(k)}\}$，而且

$$\begin{aligned}p^Tx&=p_B^Tb'-\lambda p_B^Ty^{(j)}+\lambda p_N^Te^{(j)}\\&=p^T\hat{x}+\lambda(p_j-p_B^TB^{-1}a^{(j)}).\end{aligned}$$

因为上式右端第二项为正，所以$p^Tx>p^T\hat{x}$．这与式（2.3-6）矛盾．因而$z\in\Lambda$，于是$S\subset\Lambda$．前面已经证得$\Lambda\subset S$，所以$S=\Lambda$．

（有关$B^{-1}b$有零分量的情形，即退化情形，留给读者考虑．）

定理 2.3.3 设$S=\{x\mid Ax=b,\ x\geqslant 0\}\neq\Phi$，其中$A$为$m\times n$矩阵，秩为$m$，$b\in\mathscr{R}^m$，则$S$至少有一个极方向的充分必要条件是$S$

无界.

证 若S有极方向，显然S无界（由定义可知）.

若S无极方向，则由定理2.3.2对于S中的任一点$\boldsymbol{x}$均可表示为它的极点$\boldsymbol{x}^{(1)}$，…，$\boldsymbol{x}^{(k)}$的凸组合

$$\boldsymbol{x}=\sum_{i=1}^{k}\lambda_i\boldsymbol{x}^{(i)},\ \lambda_i\geqslant 0,\quad \sum_{i=1}^{k}\lambda_i=1,$$

由此得

$$\|\boldsymbol{x}\|\leqslant\sum_{i=1}^{k}\lambda_i\|\boldsymbol{x}^{(i)}\|\leqslant\sum_{i=1}^{k}\|\boldsymbol{x}^{(i)}\|,$$

因而S有界. 所以若S无界，则S没有极方向.

2.3.3 K^{00}与$\overline{K}$的关系

定理 2.3.4 设$K\subset E$是一个非空凸锥，K^{00}为K的双极，则$K^{00}=\overline{K}$.

证 1° 设$\boldsymbol{x}\in\overline{K}$. 若$\boldsymbol{x}\in K$，则由定义可得$\boldsymbol{x}\in K^{00}$. 若$\boldsymbol{x}\notin K$，则存在点列$\{\boldsymbol{x}^n\}\subset K$，$\lim\limits_{n\to\infty}\boldsymbol{x}^n=\boldsymbol{x}$. $\forall\boldsymbol{u}\in K^0$均有$(\boldsymbol{x}^n|\boldsymbol{u})\leqslant 0$. 因为$u(\boldsymbol{x})$是连续函数，所以令$n\to\infty$得$(\boldsymbol{x}|\boldsymbol{u})\leqslant 0$，因而$\boldsymbol{x}\in K^{00}$. 由此得$\overline{K}\subset K^{00}$.

2° 因为$\overline{K}$是包含点0的凸锥（见第一章习题18），所以若$\boldsymbol{x}\notin\overline{K}$，则存在一张闭超平面严格分离$\boldsymbol{x}$与$\overline{K}$（依定理2.2.4），即存在$\boldsymbol{u}\in E'$与$\alpha\in\mathscr{R}$，使

$$(\boldsymbol{x}|\boldsymbol{u})>\alpha,\ (\overline{K}|\boldsymbol{u})<\alpha.$$

因为$0\in\overline{K}$，所以$\alpha>0$，因而$(\boldsymbol{x}|\boldsymbol{u})>0$.

因为$\overline{K}$为锥，所以$\forall\boldsymbol{y}\in\overline{K}$，$\lambda\boldsymbol{y}\in\overline{K}$（$\lambda>0$）. 而$(\lambda\boldsymbol{y}|\boldsymbol{u})=\lambda(\boldsymbol{y}|\boldsymbol{u})<\alpha$，所以$(\boldsymbol{y}|\boldsymbol{u})\leqslant 0$. 因而$\boldsymbol{u}\in K^0$. 再考虑到$(\boldsymbol{x}|\boldsymbol{u})>0$，所以$\boldsymbol{x}\notin K^{00}$. 于是得$K^{00}\subset\overline{K}$.

综上所述得$K^{00}=K$.

根据定理2.3.4，容易确定一个非空凸锥K的双极K^{00}. 例如

设$K=\{\boldsymbol{x}\mid\boldsymbol{x}\in\mathscr{R}^2,\ x_1>0,\ x_2\geqslant x_1\}$，则

$$K^{00}=\{\boldsymbol{x}\mid\boldsymbol{x}\in\mathscr{R}^2,\ x_1\geqslant 0,\ x_2\geqslant x_1\}.$$

2.3.4 多面锥与有限生成锥的关系

在§1.6中引入了有限生成的凸锥与多面锥的概念．现在证明它们是相同的．先证明以下定理．

定理 2.3.5 设T为由$\mathscr{R}^n$到$\mathscr{R}^m$的线性映射，$\boldsymbol{b}\in\mathscr{R}^m$，$C=\{\boldsymbol{x}\mid T\boldsymbol{x}\leqslant\boldsymbol{b}\}$，则

1° C至多具有限多个极点；

2° 若C非空有界，则C是一个凸多面体．

证 1° 容易验证C为闭凸集．设J为C的极点集，$\boldsymbol{p}^{(i)T}$为T的第i行．令

$$I(\boldsymbol{x})=\{i\mid\boldsymbol{p}^{(i)T}\boldsymbol{x}=b_i\}.$$

假设$\boldsymbol{x},\ \boldsymbol{y}\in J$，且$I(\boldsymbol{x})=I(\boldsymbol{y})$．因为$\forall i\overline{\in}I(\boldsymbol{x}),\boldsymbol{p}^{(i)T}\boldsymbol{x}<b_i$，所以存在$\varepsilon>0$，使

$$\boldsymbol{p}^{(i)T}(\boldsymbol{x}+\varepsilon(\boldsymbol{x}-\boldsymbol{y}))\leqslant b_i,$$

当$i\in I(\boldsymbol{x})$时，$\boldsymbol{p}^{(i)T}\boldsymbol{x}=\boldsymbol{p}^{(i)T}\boldsymbol{y}=b_i$，所以上式也成立．于是点$\boldsymbol{x}+\varepsilon(\boldsymbol{x}-\boldsymbol{y})\in C$．因为$\boldsymbol{x}$是$C$的极点，而且

$$\boldsymbol{x}=\frac{\varepsilon}{1+\varepsilon}\boldsymbol{y}+\frac{1}{1+\varepsilon}(\boldsymbol{x}+\varepsilon(\boldsymbol{x}-\boldsymbol{y})),$$

所以$\boldsymbol{x}=\boldsymbol{y}$．由此可知，若$\boldsymbol{x}\neq\boldsymbol{y}$，则$I(\boldsymbol{x})\neq I(\boldsymbol{y})$．而$\{1,\ \cdots,\ m\}$只可能具有有限多个子集$I(\boldsymbol{x})$，所以$C$只可能具有有限多个极点．

2° 因为C为非空有界闭凸集，即紧集．根据定理2.3.1 $C=\mathrm{co}(J)$，而C的极点就是它的顶点，所以C为凸多面体．

定理 2.3.6 设$K\subset\mathscr{R}^n$．K为多面锥的充分必要条件是：K为有限生成的凸锥．

证 关于必要性的证明从略，请参看文献2．下面证明充分性．

设K是有限生成的凸锥．根据定理1.6.6 K是闭的，即$\overline{K}=K$．

由定理2.3.4知$K^{00}=\overline{K}$，因而$K=K^{00}=(K^{0})^{0}$．由定理1.6.5知K^{0}是一个多面锥．由必要性得：K^{0}是一个有限生成的凸锥．再由定理1.6.5知：K^{0}的极K^{00}是一个多面锥，因而K是一个多面锥．

2.3.5 择一定理

下面要介绍的定理等价于Farkas引理．

定理 2.3.7 （**择一定理**）设A为$m\times n$矩阵，$\boldsymbol{b}\in\mathscr{R}^{m}$，则

（1） $A\boldsymbol{x}=\boldsymbol{b}$，$\boldsymbol{x}\geqslant 0$有解$\boldsymbol{x}\in\mathscr{R}^{n}$；

（2） $A^{T}\boldsymbol{y}\geqslant 0$，$\boldsymbol{b}^{T}\boldsymbol{y}<0$ 有解$\boldsymbol{y}\in\mathscr{R}^{m}$．

但不能两组都有解．

证 若(1)有解，则$\boldsymbol{b}^{T}\boldsymbol{y}=(A\boldsymbol{x})^{T}\boldsymbol{y}=\boldsymbol{x}^{T}A^{T}\boldsymbol{y}$．因为$\boldsymbol{x}\geqslant 0$，所以若$A^{T}\boldsymbol{y}\geqslant 0$，则$\boldsymbol{b}^{T}\boldsymbol{y}\geqslant 0$，因而(2)无解．

若(1)无解．令

$$K=\left\{\boldsymbol{y}\,\middle|\,\boldsymbol{y}=\sum_{j=1}^{n}\lambda_{j}\boldsymbol{a}^{(j)},\ \lambda_{j}\geqslant 0\right\}$$

其中$\boldsymbol{a}^{j}$为A的第j列，$j=1,\cdots,n$．K是有限生成的凸锥．由定理1.6.6知K是闭的．而$\boldsymbol{b}\overline{\in}K$．再由定理2.2.5知：存在包含点0的闭超平面分离$\boldsymbol{b}$与K．不妨设存在$\boldsymbol{p}\in\mathscr{R}^{m}$，$\boldsymbol{p}\neq 0$，而且$\boldsymbol{p}^{T}\boldsymbol{y}\geqslant 0$，$\forall\,\boldsymbol{y}\in K$，$\boldsymbol{p}^{T}\boldsymbol{b}<0$．因为$\boldsymbol{a}^{j}\in K$，所以$\boldsymbol{p}^{T}\boldsymbol{a}^{j}\geqslant 0$，$\forall\,j$，因而$A^{T}\boldsymbol{p}\geqslant 0$．于是$\boldsymbol{p}$是(2)的解．

2.3.6 无界整数规划的条件

本节讨论整数规划

$$(\mathrm{IP})\qquad\begin{aligned}&\max\quad z=\boldsymbol{c}^{T}\boldsymbol{x},\\&\text{s.t.}\quad A\boldsymbol{x}=\boldsymbol{b},\ \boldsymbol{x}=(x_{1},\cdots,x_{n})^{T}_{n}\geqslant 0,\\&\qquad\ x_{j}:\ \text{整数},\ j=1,\cdots,n,\end{aligned}$$

其中A为$m\times n$矩阵，秩为m，$\boldsymbol{b}\in\mathscr{R}^{m}$，$\boldsymbol{c}\in\mathscr{R}^{n}$．

整数规划是数学规划中一个重要的分支．通常是先解对应的线性规划（LP）．若（LP）不可行，则整数规划（IP）也不可

行．若（LP）无界，那么（IP）是否也无界？在一般情况下是不能肯定的．有时虽然（LP）无界，而（IP）却是有界的；有时（LP）无界，而（IP）无可行解．然而，在一定条件下答案是肯定的．

定理 2.3.8 设在（IP）中已知数据c_j，a_{ij}，b_i均为有理数，而且可行域非空．若对应的线性规划（LP）无有限的最优解（目标函数无界），则（IP）也无有限的最优解．

证 由假设可行域非空（不妨设$\boldsymbol{x}^{(0)}$为可行点）和（LP）的目标函数无界知：凸多面集

$$S=\{\boldsymbol{x}\in\mathscr{R}^n \mid A\boldsymbol{x}=\boldsymbol{b},\ \boldsymbol{x}\geqslant 0\}$$

非空而且是无界的．由定理2.3.3知，S至少有一个极方向．由定理1.7.2知，S至少有一个极点．

设$\boldsymbol{x}^{(1)}$，…，$\boldsymbol{x}^{(k)}$为S的极点，$\boldsymbol{d}^{(1)}$，…，$\boldsymbol{d}^{(l)}$为S的极方向，由定理2.3.2知：对于S中所有点$\boldsymbol{x}$均有

$$\boldsymbol{x}=\sum_{i=1}^{k}\lambda_i\boldsymbol{x}^{(i)}+\sum_{j=1}^{l}\mu_j\boldsymbol{d}^{(j)},\qquad (2.3\text{-}13)$$

其中 $\lambda_i\geqslant 0$，$i=1$，…，k，$\sum_{i=1}^{k}\lambda_i=1$，$\mu_j\geqslant 0$，$j=1$，…，$l$．

将$\boldsymbol{x}$的表达式（2.3-13）代入目标函数得

$$\boldsymbol{c}^T\boldsymbol{x}=\sum_{i=1}^{k}\lambda_i\boldsymbol{c}^T\boldsymbol{x}^{(i)}+\sum_{j=1}^{l}\mu_j\boldsymbol{c}^T\boldsymbol{d}^{(j)}.\qquad (2.3\text{-}14)$$

因为（LP）的目标函数无界，所以由式（2.3-14）知：至少有一个极方向$\boldsymbol{d}^{(j)}$，使$\boldsymbol{c}^T\boldsymbol{d}^{(j)}>0$．不妨设$\boldsymbol{c}^T\boldsymbol{d}^{(1)}>0$．再由$\boldsymbol{x}^{(0)}\in S$与极方向的定义知：$\boldsymbol{d}^{(1)}\neq 0$，而且$\boldsymbol{x}^{(0)}+\lambda\boldsymbol{d}^{(1)}\in S$，$\forall\lambda\geqslant 0$．这等价于

$$A\boldsymbol{d}^{(1)}=0,\ \boldsymbol{d}^{(1)}\geqslant 0,\ \boldsymbol{d}^{(1)}\neq 0$$

因而存在有理数$r>0$，使

$$d_1^{(1)}+d_2^{(1)}+\cdots+d_n^{(1)}\geqslant r,$$

$$\boldsymbol{c}^T\boldsymbol{d}^{(1)}\geqslant r,\ 其中\boldsymbol{d}^{(1)}=(d_1^{(1)},\ d_2^{(1)},\ \cdots,\ d_n^{(1)})^T$$

考虑线性规划：

$$(\mathrm{LP}_1)\quad \begin{aligned} \max\quad & z=\boldsymbol{c}^T\boldsymbol{y}, \\ \mathrm{s.t.}\quad & A\boldsymbol{y}=0, \\ & \boldsymbol{c}^T\boldsymbol{y}\geqslant r, \\ & y_1+y_2+\cdots+y_n\geqslant r, \\ & \boldsymbol{y}=(y_1,\ y_2,\ \cdots y_n)^T\geqslant 0. \end{aligned}$$

因为（LP_1）有可行解$\boldsymbol{d}^{(1)}$，所以必有基本可行解，可以用单纯形法求得．因为在（LP_1）的约束中已知数据均为有理数，且单纯形法仅含有理运算，所以用单纯形法得到的第一个基本可行解$\boldsymbol{y}^{(1)}$就是有理数解．用一个适当的正整数乘之，消去分母得（LP_1）的一个整数可行解，记为$\boldsymbol{y}^{(0)}$．$\boldsymbol{y}^{(0)}$也是凸多面集S的极方向．令

$$S'=\{\boldsymbol{x}\in\mathscr{R}^n\,|\,\boldsymbol{x}=\boldsymbol{x}^{(0)}+\lambda\boldsymbol{y}^{(0)},\ \lambda为非负整数\}$$

易知，无界点集S'中的点$\boldsymbol{x}$都是（IP）的可行点，因而（IP）的可行域无界．

在S'上（IP）的目标函数值为

$$z=\boldsymbol{c}^T(\boldsymbol{x}^{(0)}+\lambda\boldsymbol{y}^{(0)})=\boldsymbol{c}^T\boldsymbol{x}^{(0)}+\lambda\boldsymbol{c}^T\boldsymbol{y}^{(0)}.$$

由于$\boldsymbol{c}^T\boldsymbol{y}^0>0$，所以当$\lambda$取正整数并趋向于$+\infty$时，$z\to+\infty$，即（IP）的目标函数无界．

注　若将定理中的条件"对应的线性规划（LP）无有限的最优解"改为"在使用单纯形法的过程中发现（LP）无有限的最优解"，则证明可以简化．

习　　题

1．设$C\subset\mathscr{R}^2$为半圆：$x_1^2+x_2^2\leqslant 1$，$x_2\geqslant 0$．试验证定理2.3.1．

2．设有$\mathscr{R}^3$中四个凸集如下：

C_1：$x_1^2+(x_2+1)^2\leqslant 1$，$x_3=0$；

C_2: $(x_2-2)^2+x_3^2\leqslant 4$, $x_1=0$;

C_3: $(x_2-2)^2+x_3^2\geqslant 1$, $x_1=0$;

C_4: $x_1^2+x_2^2+x_3^2\leqslant 4$.

1° 试画出C_1，C_2的图形；

2° 是否存在分离C_1，C_2的超平面？若存在，则写出超平面的方程；该超平面是否严格分离C_1，C_2？

3° 试写出一张强分离C_1，C_3的超平面；

4° 试写出一张在点$\boldsymbol{a}=(0, 2, 0)^T$对$C_4$的支撑超平面.

3. 设$S\subset\mathscr{R}^2$由以下不等式组

$$-x_1+x_2\leqslant 0.25,\ x_1-2x_2\leqslant 4,\ x_1,\ x_2\geqslant 0$$

确定.

1° 试绘出S的图形；

2° 试写出S的全部极点；

3° 试写出S的全部极方向；

4° 试按公式（2.3-1）写出S的表达式.

*4. 设$C\subset E$为凸体，$\boldsymbol{b}\in \mathrm{bd}(C)$. 试证明存在一张在点$\boldsymbol{b}$对$C$的非平凡的闭的支撑超平面.

5. 设E是赋范线性空间，$K\subset E$为非空闭凸锥，$\boldsymbol{b}\overline{\in} K$. 试证明存在包含点0的闭超平面$H$分离$\boldsymbol{b}$与$K$.

*6. 设$C\subset\mathscr{R}^n$为凸集，$\mathrm{ri}(C)\neq\Phi$，$\boldsymbol{b}\in \mathrm{rb}(C)$. 试证明$\mathscr{R}^n$中存在一张在点$\boldsymbol{b}$对$C$的非平凡的闭的支撑超平面.

*7. 设E为赋范线性空间，$A\subset E$为非空闭凸集，$B\subset E$是非空紧凸集，而且$A\cap B=\Phi$. 试证明E中存在一张严格分离A与B的闭超平面.

8. 设C_1，$C_2\subset\mathscr{R}^n$为非空凸集. 试证明式（2.2-6）是存在超平面H严格分离C_1，C_2的充分条件，举例说明式（2.2-6）不是必要条件.

第三章　凸函数

凸函数是凸分析的主要内容之一．本章首先引入凸函数概念，然后研究凸函数的充分必要条件、凸函数的基本性质以及连续性、凸函数的闭包与凸包、凸函数的可微性、次梯度与次微分，最后研究Jensen不等式与Hadamard不等式．

§3.1　凸函数的概念

在本节中首先引入凸函数与严格凸函数的定义，然后研究$\overline{\mathscr{R}}$上的凸函数，正常凸函数与非正常凸函数，最后研究凸函数的上图．

3.1.1　凸函数与严格凸函数

定义 3.1.1　设$C\subset V$为凸集，$f:C\to\mathscr{R}$．若$\forall\, \boldsymbol{x}$，$\boldsymbol{y}\in C$，$\lambda\in(0,1)$均有

$$f(\lambda\boldsymbol{x}+(1-\lambda)\boldsymbol{y})\leqslant\lambda f(\boldsymbol{x})+(1-\lambda)f(\boldsymbol{y}),\qquad(3.1\text{-}1)$$

则称$f(\boldsymbol{x})$为凸集C上的**凸函数**（convex function）．

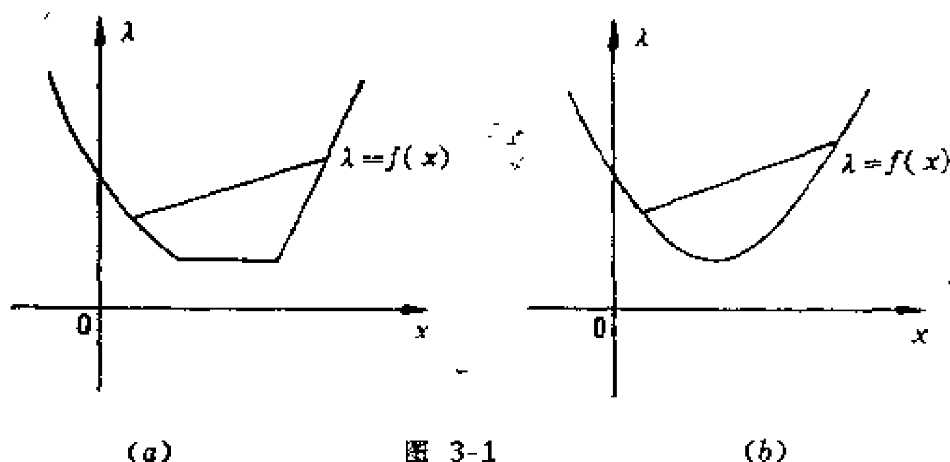
定义 3.1.2　设$C\subset V$为凸集，$f:C\to\mathscr{R}$．若$\forall\, \boldsymbol{x}$，$\boldsymbol{y}\in C$，$\boldsymbol{x}\neq\boldsymbol{y}$，$\lambda\in(0,1)$都有

$$f(\lambda\boldsymbol{x}+(1-\lambda)\boldsymbol{y})<\lambda f(\boldsymbol{x})+(1-\lambda)f(\boldsymbol{y}),\qquad(3.1\text{-}2)$$

则称$f(\boldsymbol{x})$为凸集C上的**严格凸函数**．

若式(3.1-1)中的不等号为“$\geqslant$”，则称$f(\boldsymbol{x})$为凸集C上的凹函数．若式(3.1-2)中的不等号为“$>$”，则$f(\boldsymbol{x})$称为凸集C上的严格凹函数．

凸函数的**几何意义**是：连接函数图形上任意两点的线段，处处都不在函数图形的下方（图3-1(a)）．严格凸函数的几何意义是：



(a) 图 3-1 (b)

连接函数图形上任意两点的线段处处都在函数图形的上方（图 3-1 (b)），两端除外。

显然，严格凸函数是凸函数，严格凹函数也是凹函数。

定义 3.1.3 设$C\subset\mathscr{R}^n$为凸集，$f:C\to\mathscr{R}$．若存在常数$k>0$，使$\forall\, \boldsymbol{x},\ \boldsymbol{y}\in C,\ \lambda\in(0,\ 1)$均有

$$f(\lambda\boldsymbol{x}+(1-\lambda)\boldsymbol{y})$$
$$\leqslant\lambda f(\boldsymbol{x})+(1-\lambda)f(\boldsymbol{y})-k\lambda(1-\lambda)\|\boldsymbol{x}-\boldsymbol{y}\|^2, \qquad (3.1\text{-}3)$$

则称$f(\boldsymbol{x})$在凸集C上是一致凸的。

定义 3.1.4 设$f:V\to\mathscr{R}$为线性函数，$a\in\mathscr{R}$．则$F(\boldsymbol{x})=f(\boldsymbol{x})+a$称为**仿射函数**(affine fuction)。

例 1 设$F(\boldsymbol{x})$为V上的仿射函数．试证明$F(\boldsymbol{x})$既是凸函数，又是凹函数。

证 设$\boldsymbol{x},\ \boldsymbol{y}\in V$，$0<\lambda<1$．根据定义得

$$F(\lambda\boldsymbol{x}+(1-\lambda)\boldsymbol{y})=f(\lambda\boldsymbol{x}+(1-\lambda)\boldsymbol{y})+a$$
$$=\lambda f(\boldsymbol{x})+(1-\lambda)f(\boldsymbol{y})+a$$
$$=\lambda[f(\boldsymbol{x})+a]+(1-\lambda)[f(\boldsymbol{y})+a]$$
$$=\lambda F(\boldsymbol{x})+(1-\lambda)F(\boldsymbol{y})\leqslant\lambda F(\boldsymbol{x})+(1-\lambda)F(\boldsymbol{y}),$$

所以$F(\boldsymbol{x})$是凸函数．同理可证$F(\boldsymbol{x})$是凹函数。

例 2 设$f(x)=(x-1)^2$，$x\in\mathscr{R}$．试证明$f(x)$在区间$(-\infty,$

$+\infty$）内是严格凸函数.

证 设x，y为任意二实数，$x\neq y$，$\lambda\in(0,1)$. 因为

$$
\begin{aligned}
& f(\lambda x+(1-\lambda)y) \\
&=[\lambda x+(1-\lambda)y]^2-2[\lambda x+(1-\lambda)y]+1 \\
&=[y+\lambda(x-y)][x+(1-\lambda)(y-x)] \\
&\quad -2[\lambda x+(1-\lambda)y]+1 \\
&=\lambda x^2+(1-\lambda)y^2-\lambda(1-\lambda)(y-x)^2-2\lambda x-2(1-\lambda)y+1 \\
&=\lambda(x^2-2x+1)+(1-\lambda)(y^2-2y+1)-\lambda(1-\lambda)(y-x)^2 \\
&<\lambda(x^2-2x+1)+(1-\lambda)(y^2-2y+1) \\
&=\lambda f(x)+(1-\lambda)f(y),
\end{aligned}
$$

所以

$$f(\lambda x+(1-\lambda)y)<\lambda f(x)+(1-\lambda)f(y).$$

由x，y的任意性与λ的任意性知：$f(x)$在区间（$-\infty$，$+\infty$）内是严格凸函数.

3.1.2 在$\overline{\mathscr{R}}$中取值的凸函数

下面讨论在$\overline{\mathscr{R}}=\{-\infty\}\cup\mathscr{R}\cup\{+\infty\}$中取值的凸函数. 因为可能涉及到包含$+\infty$或$-\infty$的运算，所以在本书中作以下规定.

$$
\begin{aligned}
&(+\infty)+a=+\infty, && a\neq-\infty \\
&(-\infty)+a=-\infty, && a\neq+\infty \\
&(+\infty)\cdot a=+\infty,\ (-\infty)\cdot a=-\infty, && a>0 \\
&(+\infty)\cdot a=-\infty,\ (-\infty)\cdot a=+\infty, && a<0 \\
&0\cdot(+\infty)=0\cdot(-\infty)=0,\ -(-\infty)=+\infty, \\
&\inf\Phi=+\infty,\ \sup\Phi=-\infty.
\end{aligned}
$$

$(+\infty)+(-\infty)$，$(+\infty)-(+\infty)$与$(-\infty)-(-\infty)$都无意义，应当避免.

定义 3.1.5 设$C\subset V$为凸集，$f:C\to\overline{\mathscr{R}}$. 若对于所有的$x$，$y\in C$与所有满足$f(x)<\mu$，$f(y)<\nu$的$\mu$，$\nu\in\mathscr{R}$，当$0<\lambda<1$时均有

$$f(\lambda \boldsymbol{x}+(1-\lambda)\boldsymbol{y})<\lambda\mu+(1-\lambda)\nu, \qquad (3.1\text{-}4)$$

则称$f(\boldsymbol{x})$为**凸函数**.

定理 3.1.1 设$C\subset V$为凸集，$f:C\to\mathscr{R}$. 则定义3.1.5与定义3.1.1等价.

证 设按定义3.1.1 $f(\boldsymbol{x})$是凸函数，且μ，$\nu\in\mathscr{R}$，$f(\boldsymbol{x})<\mu$，$f(\boldsymbol{y})<\nu$，$0<\lambda<1$. 于是

$$\begin{aligned} f(\lambda \boldsymbol{x}+(1-\lambda)\boldsymbol{y}) &\leqslant \lambda f(\boldsymbol{x})+(1-\lambda)f(\boldsymbol{y}) \\ &< \lambda\mu+(1-\lambda)\nu, \end{aligned}$$

因而按定义3.1.5 $f(\boldsymbol{x})$也是凸函数.

设按义3.1.5 $f(\boldsymbol{x})$是凸函数. 因为$\forall\varepsilon>0$均有$f(\boldsymbol{x})<f(\boldsymbol{x})+\varepsilon$，$f(\boldsymbol{y})<f(\boldsymbol{y})+\varepsilon$，所以

$$\begin{aligned} f(\lambda \boldsymbol{x}+(1-\lambda)\boldsymbol{y}) &< \lambda[f(\boldsymbol{x})+\varepsilon]+(1-\lambda)[f(\boldsymbol{y})+\varepsilon] \\ &= \lambda f(\boldsymbol{x})+(1-\lambda)f(\boldsymbol{y})+\varepsilon. \end{aligned}$$

令$\varepsilon\to 0$得

$$f(\lambda \boldsymbol{x}+(1-\lambda)\boldsymbol{y})\leqslant\lambda f(\boldsymbol{x})+(1-\lambda)f(\boldsymbol{y}),$$

因而按定义3.1.1 $f(\boldsymbol{x})$也是凸函数.

定义 3.1.6 设$C\subset V$为凸集，$f:C\to\overline{\mathscr{R}}$为凸函数，则称集合$\{\boldsymbol{x}\mid \boldsymbol{x}\in C,\ f(\boldsymbol{x})<+\infty\}$为$f(\boldsymbol{x})$的**有效域**（effective domain），记为dom(f).

性质 凸函数$f(x)$的有效域dom(f)是凸集（证明留给读者）.

定义 3.1.7 设$f:V\to\mathscr{R}\cup\{+\infty\}$为凸函数，而且dom$(f)\neq\Phi$，则称$f(\boldsymbol{x})$为**正常凸函数**（proper convex function）.

定义 3.1.8 若$f(\boldsymbol{x})$为V上的凸函数，但不是正常凸函数，则称$f(\boldsymbol{x})$为非正常凸函数.

注 定义在凸集$C\subset V$上的实值凸函数$f(\boldsymbol{x})$可以开拓为整个空间V上的正常凸函数：

$$F(\boldsymbol{x})=\begin{cases} f(\boldsymbol{x}), & x\in C; \\ +\infty, & x\overline{\in} C, \end{cases} \qquad (3.1\text{-}5)$$

而且有效域不变：$\mathrm{dof}(F)=\mathrm{doof}(f)=C$.

反之，每个正常凸函数 $f(\boldsymbol{x})$ 都可以看作定义在它的有效域 $\mathrm{dom}(f)$上的实值凸函数的开拓．注意到这一点将会带来许多方便．

正常凹函数的定义与正常凸函数的定义类似，从略，不再赘述．

3.1.3 凸函数的上图

由图3-1可以直观地看出：位于凸函数$f(\boldsymbol{x})$的图形上方的集合是个凸集．有时利用这个集合的性质研究函数$f(\boldsymbol{x})$的性质更方便，所以引入以下定义，并研究函数$f(\boldsymbol{x})$的凸性与上述集合的凸性之间的关系．

定义 3.1.9 设$C\subset V$，函数$f:C\to\overline{\mathscr{R}}$．则称集合

$$\{(\boldsymbol{x},\ \lambda)\mid(\boldsymbol{x},\ \lambda)\in C\times\mathscr{R},\ f(\boldsymbol{x})\leqslant\lambda\} \tag{3.1-6}$$

为$f(\boldsymbol{x})$的上图（epigraph），记为$\mathrm{epi}(f)$.

例 3 设函数$f(x)$定义如下：

$$f(x)=\begin{cases}x^2+1, & -\infty<x\leqslant-1;\\ x^2, & -1<x\leqslant 1;\\ x+1, & 1<x<+\infty.\end{cases} \tag{3.1-7}$$

试描绘$f(x)$的上图$\mathrm{epi}(f)$．再判断它的闭性．

解 根据定义可描绘出$f(x)$的上图$\mathrm{epi}(f)$如图3-2所示．

设 $\boldsymbol{a}=(1,\ 1)^T$，$\boldsymbol{b}=(1,\ 2)^T$，

$\boldsymbol{c}=(-1,\ 2)^T$，$\boldsymbol{d}=(-1,\ 1)^T$.

则 $[\boldsymbol{a},\ \boldsymbol{b}]\subset\mathrm{epi}(f)$，$(\boldsymbol{c},\ \boldsymbol{d}]\overline{\subset}\mathrm{epi}(f)$.

所以$\mathrm{epi}(f)$不是闭集．

由图3-2可以看出，函数$f(\boldsymbol{x})$的上图$\mathrm{epi}(f)$不一定是凸集，也不一定是闭集．然而，凸函数的上图必定是凸集；反之，上图为凸集的函数必为凸函数．这种对应关系将在定理3.1.2中描述．关于

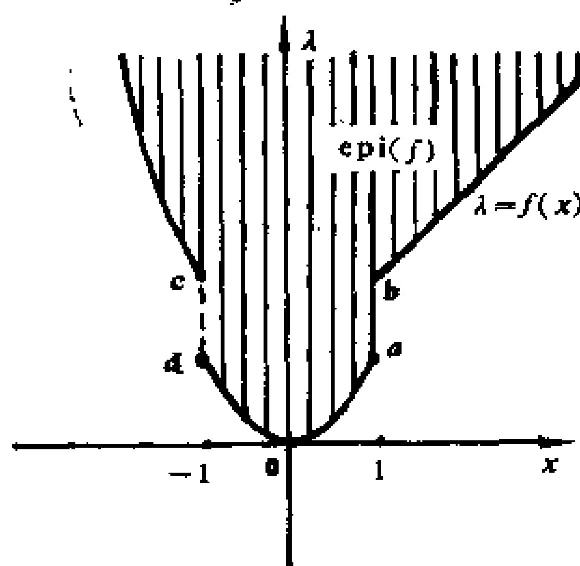


图 3-2

函数的上图为闭集的等价条件将在§3.4中讨论.

有时也用到$f(\boldsymbol{x})$的下图，也就是集合

$$\{(\boldsymbol{x}, \lambda) \mid (\boldsymbol{x}, \lambda) \in C \times \mathscr{R}, \ \lambda \leqslant f(\boldsymbol{x})\}.$$

定理 3.1.2 设$f: V \to \overline{\mathscr{R}}$，则下列条件等价：

1° f是凸函数；

2° $\mathrm{epi}(f)$是凸集；

3° $A=\{(\boldsymbol{x}, \lambda) \mid (\boldsymbol{x}, \lambda) \in V \times \mathscr{R}, \ f(\boldsymbol{x})<\lambda\}$是凸集.

证 "1°⇒2°" 设$f(\boldsymbol{x})$是凸函数. 再设$(\boldsymbol{x}, \mu)$，$(\boldsymbol{y}, \nu) \in \mathrm{epi}(f)$，$0<\lambda<1$.

因为$(\boldsymbol{x}, \mu)$，$(\boldsymbol{y}, \nu) \in \mathrm{epi}(f)$，所以$f(\boldsymbol{x}) \leqslant \mu$，$f(\boldsymbol{y}) \leqslant \nu$. 因而$f(\boldsymbol{x})<\mu+\varepsilon$，$f(\boldsymbol{y})<\nu+\varepsilon$，$\forall \varepsilon>0$. 再根据$f(\boldsymbol{x})$的凸性得

$$\begin{aligned} f(\lambda \boldsymbol{x}+(1-\lambda) \boldsymbol{y}) &< \lambda(\mu+\varepsilon)+(1-\lambda)(\nu+\varepsilon) \\ &= \lambda \mu+(1-\lambda) \nu+\varepsilon. \end{aligned}$$

令$\varepsilon \to 0$得

$$f(\lambda \boldsymbol{x}+(1-\lambda) \boldsymbol{y}) \leqslant \lambda \mu+(1-\lambda) \nu.$$

于是 $\lambda(\boldsymbol{x}, \mu)+(1-\lambda)(\boldsymbol{y}, \nu)$

$$=(\lambda \boldsymbol{x}+(1-\lambda)\boldsymbol{y},\ \lambda\mu+(1-\lambda)\nu)\in \mathrm{epi}(f),$$

所以 epi(f)是凸集。

"$2°\Rightarrow 3°$" 设epi(f)是凸集. 再设$(\boldsymbol{x},\ \mu)$, $(\boldsymbol{y},\ \nu)\in A$, $0<\lambda<1$.

因为$(\boldsymbol{x},\ \mu)$, $(\boldsymbol{y},\ \nu)\in A$, 所以$f(\boldsymbol{x})<\mu, f(\boldsymbol{y})<\nu$. 存在$\mu_0$, $\nu_0\in\mathscr{R}$满足$f(\boldsymbol{x})\leqslant\mu_0<\mu$, $f(\boldsymbol{y})\leqslant\nu_0<\nu$. 因而$(\boldsymbol{x},\ \mu_0)$, $(\boldsymbol{y},\ \nu_0)\in \mathrm{epi}(f)$. 因为epi($f$)是凸集, 所以

$$\lambda(\boldsymbol{x},\ \mu_0)+(1-\lambda)(\boldsymbol{y},\ \nu_0)$$
$$=(\lambda \boldsymbol{x}+(1-\lambda)\boldsymbol{y},\ \lambda\mu_0+(1-\lambda)\nu_0)\in \mathrm{epi}(f),$$

于是

$$f(\lambda \boldsymbol{x}+(1-\lambda)\boldsymbol{y})\leqslant\lambda\mu_0+(1-\lambda)\nu_0<\lambda\mu+(1-\lambda)\nu.$$

因而

$$\lambda(\boldsymbol{x},\ \mu)+(1-\lambda)(\boldsymbol{y},\ \nu)$$
$$=(\lambda \boldsymbol{x}+(1-\lambda)\boldsymbol{y},\ \lambda\mu+(1-\lambda)\nu)\in A,$$

所以A是凸集.

"$3°\Rightarrow 1°$" 设A是凸集. 再设$f(\boldsymbol{x})<\mu$, $f(\boldsymbol{y})<\nu$, $0<\lambda<1$. 于是$(\boldsymbol{x},\ \mu)$, $(\boldsymbol{y},\ \nu)\in A$. 由$A$的凸性得

$$\lambda(\boldsymbol{x},\ \mu)+(1-\lambda)(\boldsymbol{y},\ \nu)$$
$$=(\lambda \boldsymbol{x}+(1-\lambda)\boldsymbol{y},\ \lambda\mu+(1-\lambda)\mu)\in A.$$

于是

$$f(\lambda \boldsymbol{x}+(1-\lambda)\boldsymbol{y})<\lambda\mu+(1-\lambda)\nu,$$

因而$f(\boldsymbol{x})$为凸函数.

因为函数$f(\boldsymbol{x})$的凸性和它的上图epi(f)的凸性等价, 所以也可以把"epi(f)是凸集"作为凸函数的定义, 而把前面的定义作为凸函数的性质. 要证明$f(\boldsymbol{x})$为凸函数, 也可以先证明$f(\boldsymbol{x})$的上图epi(f)为凸集, 或集合$A=\{(\boldsymbol{x},\ \lambda)\mid f(\boldsymbol{x})<\lambda\}$为凸集.

例 4 设$B\subset V\times\mathscr{R}$是凸集. 函数$f:V\to\overline{\mathscr{R}}$定义如下:

$$f(\boldsymbol{x})=\inf\{\lambda\mid(\boldsymbol{x},\ \lambda)\in B\}. \qquad (3.1\text{-}8)$$

试证明$f(\boldsymbol{x})$为凸函数（注：$\inf\Phi=+\infty$），且$\text{epi}(f)\supset B$.

证 设$(\boldsymbol{x}, \lambda_1)$，$(\boldsymbol{y}, \lambda_2)\in B$，且$f(\boldsymbol{x})<\lambda_1$，$f(\boldsymbol{y})<\lambda_2$. 由$f(\boldsymbol{x})$的定义可知：存在$\mu<\lambda_1$，$\nu<\lambda_2$，使$(\boldsymbol{x}, \mu)$，$(\boldsymbol{y}, \nu)\in B$. 因为$B$是凸集，所以$\forall\,\alpha\in(0, 1)$均有

$$(\alpha\boldsymbol{x}+(1-\alpha)\boldsymbol{y},\ \alpha\mu+(1-\alpha)\nu)$$
$$=\alpha(\boldsymbol{x}, \mu)+(1-\alpha)(\boldsymbol{y}, \nu)\in B,$$

因而

$$f(\alpha\boldsymbol{x}+(1-\alpha)\boldsymbol{y})\leqslant\alpha\mu+(1-\alpha)\nu<\alpha\lambda_1+(1-\alpha)\lambda_2,$$

所以$f(\boldsymbol{x})$是凸函数.

§3.2 凸函数的充分必要条件

按照凸函数的定义检查函数的凸性是比较困难的. 为了便于应用，下面提出几个关于凸函数的充分必要条件，供选择使用.

定理 3.2.1 设$f: V\to\overline{\mathscr{R}}$. 则$f(\boldsymbol{x})$为凸函数的充分必要条件是：$f(\boldsymbol{x})$在$V$中的每条直线$m$上的限制（记为$f|m$）是凸的.

证 由定义3.1.5直接可得.

定理 3.2.2 设在开凸集$D\subset\mathscr{R}^n$内$f(\boldsymbol{x})\in C^1$，则$f(\boldsymbol{x})$为凸函数的充分必要条件为：$\forall\,\boldsymbol{x}, \boldsymbol{y}\in D$都有

$$f(\boldsymbol{y})\geqslant f(\boldsymbol{x})+\nabla f(\boldsymbol{x})^T(\boldsymbol{y}-\boldsymbol{x}), \tag{3.2-1}$$

其中

$$\nabla f(\boldsymbol{x})^T=\left(\frac{\partial f}{\partial x_1}, \cdots, \frac{\partial f}{\partial x_n}\right).$$

证 必要性. 设$f(\boldsymbol{x})$为凸函数. 任取$\boldsymbol{x}$，$\boldsymbol{y}\in D$，$\alpha\in(0, 1)$. 由$f(\boldsymbol{x})$的凸性得

$$f(\alpha\boldsymbol{y}+(1-\alpha)\boldsymbol{x})\leqslant\alpha f(\boldsymbol{y})+(1-\alpha)f(\boldsymbol{x}),$$

即

$$f(\boldsymbol{x}+\alpha(\boldsymbol{y}-\boldsymbol{x}))\leqslant f(\boldsymbol{x})+\alpha(f(\boldsymbol{y})-f(\boldsymbol{x})). \tag{3.2-2}$$

根据泰勒公式得

$$f(\boldsymbol{x}+\alpha(\boldsymbol{y}-\boldsymbol{x}))=f(\boldsymbol{x})+\alpha\nabla f(\boldsymbol{\xi})^T(\boldsymbol{y}-\boldsymbol{x}), \qquad (3.2-3)$$

其中$\boldsymbol{\xi}=\boldsymbol{x}+\theta\alpha(\boldsymbol{y}-\boldsymbol{x})$，$0<\theta<1$．由于$D$为凸集，所以$\boldsymbol{\xi}\in D$．由式(3.2-2)与(3.2-3)得

$$f(\boldsymbol{y})-f(\boldsymbol{x})\geqslant\nabla f(\boldsymbol{\xi})^T(\boldsymbol{y}-\boldsymbol{x}),$$

令$\theta\to0$，则$\nabla f(\boldsymbol{\xi})\to\nabla f(\boldsymbol{x})$．因而上式化为式(3.2-1)．

充分性．设$\forall\boldsymbol{x}$，$\boldsymbol{y}\in D$式(3.2-1)都成立．再设$\boldsymbol{x}$，$\boldsymbol{y}\in D$，$\alpha\in(0,1)$，$\boldsymbol{z}=\alpha\boldsymbol{x}+(1-\alpha)\boldsymbol{y}$．由$D$的凸性得$\boldsymbol{z}\in D$．由假设得

$$f(\boldsymbol{x})\geqslant f(\boldsymbol{z})+\nabla f(\boldsymbol{z})^T(\boldsymbol{x}-\boldsymbol{z}),$$

$$f(\boldsymbol{y})\geqslant f(\boldsymbol{z})+\nabla f(\boldsymbol{z})^T(\boldsymbol{y}-\boldsymbol{z}).$$

用α，$1-\alpha$分别乘上面二式两端，再相加得

$$\begin{aligned}\alpha f(\boldsymbol{x})+(1-\alpha)f(\boldsymbol{y})&\geqslant f(\boldsymbol{z})+\nabla f(\boldsymbol{z})^T(\alpha\boldsymbol{x}+(1-\alpha)\boldsymbol{y}-\boldsymbol{z})\\&=f(\alpha\boldsymbol{x}+(1-\alpha)\boldsymbol{y}).\end{aligned}$$

于是

$$f(\alpha\boldsymbol{x}+(1-\alpha)\boldsymbol{y})\leqslant\alpha f(\boldsymbol{x})+(1-\alpha)f(\boldsymbol{y}),$$

因而$f(\boldsymbol{x})$是凸函数．

定理 3.2.3 设在开凸集$D\subset\mathscr{R}^n$内$f(\boldsymbol{x})\in C^1$，则$f(\boldsymbol{x})$为严格凸函数的充分必要条件是：$\forall\boldsymbol{x}$，$\boldsymbol{y}\in D$，$\boldsymbol{x}\neq\boldsymbol{y}$，都有

$$f(\boldsymbol{y})>f(\boldsymbol{x})+\nabla f(\boldsymbol{x})^T(\boldsymbol{y}-\boldsymbol{x}). \qquad (3.2-4)$$

证 必要性．设$f(\boldsymbol{x})$在D内为严格凸函数，再设存在两个点$\boldsymbol{x}$，$\boldsymbol{y}\in D$，$\boldsymbol{x}\neq\boldsymbol{y}$，使严格不等式(3.2-4)不成立．

因为严格凸函数也是凸函数，所以根据假设与定理3.2.2得

$$f(\boldsymbol{y})=f(\boldsymbol{x})+\nabla f(\boldsymbol{x})^T(\boldsymbol{y}-\boldsymbol{x}). \qquad (3.2-5)$$

由$f(\boldsymbol{x})$的严格凸性得

$$f\left(\frac{\boldsymbol{x}+\boldsymbol{y}}{2}\right)<\frac{1}{2}f(\boldsymbol{x})+\frac{1}{2}f(\boldsymbol{y}). \qquad (3.2-6)$$

将式(3.2-5)代入式(3.2-6)得

$$f\left(\frac{x+y}{2}\right)<f(x)+\frac{1}{2}\nabla f(x)^T(y-x). \quad (3.2\text{-}7)$$

然而根据定理3.2.2应有

$$f\left(\frac{x+y}{2}\right)\geqslant f(x)+\nabla f(x)^T\left(\frac{x+y}{2}-x\right)$$

$$=f(x)+\frac{1}{2}\nabla f(x)^T(y-x).$$

上式与式(3.2-7)矛盾，所以$\forall x, y\in D, x\neq y$，不等式(3.2-4)都成立.

充分性的证明与定理3.2.2的证明类似，从略.

定理 3.2.4 设$D\subset\mathscr{R}^n$为非空开凸集，$f: D\to\mathscr{R}$，$f(x)\in C^2$. 则$f(x)$为凸函数的充分必要条件是：在D内每一点x，$f(x)$的Hesse矩阵$H(x)$都是半正定的.

证 必要性. 设$f(x)$为凸函数，$x\in D$，$y\in\mathscr{R}^n$.

因为D是开集，所以当$t>0$充分小时，$x+ty\in D$. 根据定理3.2.2得：

$$f(x+ty)\geqslant f(x)+t\nabla f(x)^T y. \quad (3.2\text{-}8)$$

根据泰勒公式得

$$f(x+ty)=f(x)+t\nabla f(x)^T y+\frac{t^2}{2}y^T H(\xi)y \quad (3.2\text{-}9)$$

其中$\xi=x+\theta ty$，$0<\theta<1$. 因为$t^2>0$，所以由式(3.2-8)与式(3.2-9)得$y^T H(\xi)y\geqslant 0$. 令$t\to 0$，于是得$y^T H(x)y\geqslant 0$. 所以$H(x)$是半正定的.

充分性. 设在D内每一点x，$H(x)$都是半正定的. 再设$x, y\in D$. 根据泰勒公式得

$$f(y)=f(x)+\nabla f(x)^T(y-x)+\frac{1}{2}(y-x)^T H(\xi)(y-x),$$

其中$\xi=x+\theta(y-x)$，$0<\theta<1$. 因为$\xi\in D$，所以$H(\xi)$是半正

定的，因而

$$(\boldsymbol{y}-\boldsymbol{x})^T H(\xi)(\boldsymbol{y}-\boldsymbol{x})\geqslant 0,$$

于是得

$$f(\boldsymbol{y})\geqslant f(\boldsymbol{x})+\nabla f(\boldsymbol{x})^T(\boldsymbol{y}-\boldsymbol{x}),\quad \forall \boldsymbol{x},\ \boldsymbol{y}\in D.$$

根据定理3.2.2 $f(\boldsymbol{x})$是凸函数.

推论 设单元函数$f(x)$在区间$(a,\ b)$内具有二阶导数，则$f(x)$为凸函数的充分必要条件是：$f''(x)\geqslant 0$.

（因为单元函数$f(x)$的Hesse矩阵半正定等价于$f''(x)\geqslant 0$）

定理 3.2.5 设$D\subset\mathscr{R}^n$为非空开凸集，$f:D\to\mathscr{R}$，$f(\boldsymbol{x})\in C^2$. 若在D内每一点$\boldsymbol{x}$，$f(\boldsymbol{x})$的Hesse矩阵$H(\boldsymbol{x})$都是正定的，则$f(\boldsymbol{x})$为严格凸函数.

证 设$\boldsymbol{x}$，$\boldsymbol{y}\in D$，$\boldsymbol{x}\neq\boldsymbol{y}$. 根据泰勒公式得

$$f(\boldsymbol{y})=f(\boldsymbol{x})+\nabla f(\boldsymbol{x})^T(\boldsymbol{y}-\boldsymbol{x})+\frac{1}{2}(\boldsymbol{y}-\boldsymbol{x})^T H(\xi)(\boldsymbol{y}-\boldsymbol{x})$$

其中$\xi=\boldsymbol{x}+\theta(\boldsymbol{y}-\boldsymbol{x})$，$0<\theta<1$. 因为$\xi\in D$，所以$H(\xi)$是正定的. 考虑到$\boldsymbol{x}\neq\boldsymbol{y}$，于是由上式得

$$f(\boldsymbol{y})>f(\boldsymbol{x})+\nabla f(\boldsymbol{x})^T(\boldsymbol{y}-\boldsymbol{x}).$$

根据定理3.2.3 $f(\boldsymbol{x})$是严格凸函数.

推论 设单元函数$f(x)$在区间$(a,\ b)$内具有二阶导数. 若$f''(x)>0$，$\forall x\in(a,\ b)$，则$f(\boldsymbol{x})$为严格凸函数.

注 定理3.2.5的逆命题不成立. 例如$f(x)=x^4$是严格凸函数（可以证明：若$f'(x)$单调增，则$f(x)$是严格凸函数），而它的Hesse矩阵$f''(x)=12x^2$在$x=0$处为零，不是正定的.

定理 3.2.6 设$D\subset\mathscr{R}^n$为非空开凸集，$f:D\to\mathscr{R}$，$f(\boldsymbol{x})\in C^1$. $f(\boldsymbol{x})$在D上是一致凸的充分必要条件是：存在常数$k>0$，使$\forall\boldsymbol{x}$，$\boldsymbol{y}\in D$均有

$$f(\boldsymbol{y})\geqslant f(\boldsymbol{x})+\nabla f(\boldsymbol{x})^T(\boldsymbol{y}-\boldsymbol{x})+k\|\boldsymbol{y}-\boldsymbol{x}\|^2. \quad (3.2\text{-}10)$$

例 1 设$G=(g_{ij})$为$n\times n$对称半正定矩阵，$\boldsymbol{c}=(c_1,\cdots,c_n)^T$.

试证明二次函数

$$f(\boldsymbol{x})=\frac{1}{2}\boldsymbol{x}^T G\boldsymbol{x}+\boldsymbol{c}^T\boldsymbol{x} \qquad (3.2\text{-}11)$$

为凸函数.

证
$$f(\boldsymbol{x})=\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}g_{ij}x_i x_j+\sum_{j=1}^{n}c_j x_j .$$

对x_k求导得

$$f_k=\frac{1}{2}\left(\sum_{j=1}^{n}g_{kj}x_j+\sum_{i=1}^{n}g_{ik}x_i\right)+c_k$$

$$=\sum_{j=1}^{n}g_{kj}x_j+c_k,$$

再对x_l求导得$f_{kl}=g_{kl}$，因而$f(\boldsymbol{x})$的Hesse矩阵$H(\boldsymbol{x})=G$. 而G是半正定矩阵；所以$f(\boldsymbol{x})$为凸函数.

（**注**：若G为正定矩阵，则二次函数(3.2-11)为严格凸函数）

例 2 设$n\in\mathcal{N}$，x，$y>0$. 试证明：当$n>1$时恒有

$$\left(\frac{x+y}{2}\right)^n<\frac{1}{2}(x^n+y^n). \qquad (3.2\text{-}12)$$

证 令$f(t)=t^n$，$t>0$，则$f''=n(n-1)t^{n-2}>0$.所以$f(t)$是严格凸函数，依定义式(3.2-12)恒成立.

例 3 设$f_1(x)=e^x$，$f_2(x)=e^{-x}$，$x\in\mathscr{R}$. 试证明集合

$$A=\{(x,\lambda)\mid f_1(x)\leqslant\lambda,\ f_2(x)\leqslant\lambda\}$$

为凸集.

证 因为$f_1''(x)=e^x>0$，所以$f_1(x)$为凸函数，根据定理3.1.2 $f_1(x)$的上图epi(f_1)为凸集. 同理$f_2(x)$的上图epi(f_2)也是凸集. 而

$$A=\text{epi}(f_1)\cap\text{epi}(f_2),$$

所以A为凸集.

§3.3 凸函数的基本性质

性质1 设f_1，f_2为凸集$C\subset V$上的凸函数，k_1，$k_2\geqslant 0$，则$k_1f_1+k_2f_2$也是C上的凸函数.

性质2 设$f:V\to\overline{\mathscr{R}}$为凸函数，$\beta\in\mathscr{R}$. 则水平集

$$L=(f,\ \beta)=\{\boldsymbol{x}\mid f(\boldsymbol{x})\leqslant\beta\} \tag{3.3-1}$$

为凸集.

证 设$\boldsymbol{x}$，$\boldsymbol{y}\in L(f,\ \beta)$，$\lambda\in(0,\ 1)$. 再令$\varepsilon>0$.于是有

$$f(\boldsymbol{x})\leqslant\beta<\beta+\varepsilon,\ f(\boldsymbol{y})\leqslant\beta<\beta+\varepsilon.$$

根据定义3.1.5得:

$$f(\lambda\boldsymbol{x}+(1-\lambda)\boldsymbol{y})<\lambda(\beta+\varepsilon)+(1-\lambda)(\beta+\varepsilon)=\beta+\varepsilon.$$

令$\varepsilon\to 0$得$f(\lambda\boldsymbol{x}+(1-\lambda)\boldsymbol{y})\leqslant\beta$，因而$\lambda\boldsymbol{x}+(1-\lambda)\boldsymbol{y}\in L(f,\ \beta)$.所以$L(f,\ \beta)$是凸集.

性质3 设$f_k:V\to\overline{\mathscr{R}}$为凸函数序列，$k=1,\ 2,\ \cdots$，$f(\boldsymbol{x})$为$f_k(\boldsymbol{x})$的点态极限（pointwise limit），即

$$\lim_{k\to\infty}f_k(\boldsymbol{x})=f(\boldsymbol{x}), \tag{3.3-2}$$

则$f(\boldsymbol{x})$为凸函数.

证 设$\boldsymbol{x}$，$\boldsymbol{y}\in V$，μ，ν，$\lambda\in\mathscr{R}$，使

$$f(\boldsymbol{x})<\mu,\ f(\boldsymbol{y})<\nu,\ 0<\lambda<1.$$

存在$\varepsilon>0$使$f(\boldsymbol{x})<\mu-\varepsilon$，$f(\boldsymbol{y})<\nu-\varepsilon$. 由式(3.3－2)知，存在$K$，当$k>K$时，$f_k(\boldsymbol{x})<\mu-\varepsilon$，$f_k(\boldsymbol{y})<\nu-\varepsilon$. 根据$f_k(\boldsymbol{x})$的凸性得

$$f_k(\lambda\boldsymbol{x}+(1-\lambda)\boldsymbol{y})<\lambda\mu+\lambda\nu-\varepsilon,\ \forall k>K,$$

因而

$$f(\lambda\boldsymbol{x}+(1-\lambda)\boldsymbol{y})\leqslant\lambda\mu+\lambda\nu-\varepsilon$$
$$<\lambda\mu+\lambda\nu,$$

所以$f(\boldsymbol{x})$为凸函数.

性质4 设 $f_i: V\to\overline{\mathscr{R}}$ 为凸函数，其中 $i\in I$，而 I 为任一指标集. $f(\boldsymbol{x})$ 为 $f_i(\boldsymbol{x})$ 的点态上确界

$$f(\boldsymbol{x})=\sup_{i\in I}\{f_i(\boldsymbol{x})\}, \tag{3.3-3}$$

则 $f(\boldsymbol{x})$ 是凸函数.

证 因为 $f_i(\boldsymbol{x})$ 为凸函数，所以上图 $\mathrm{epi}(f_i)$ 为凸集. 要证明 $f(\boldsymbol{x})$ 为凸函数，只需证明

$$\mathrm{epi}(f)=\{\boldsymbol{x},\ \lambda)\mid(\boldsymbol{x},\ \lambda)\in V\times\mathscr{R},\ f(\boldsymbol{x})\leqslant\lambda\}$$

为凸集，而 $f(x)\leqslant\lambda$，即

$$\sup_{i\in I}\{f_i(\boldsymbol{x})\}\leqslant\lambda,$$

它等价于 $f_i(\boldsymbol{x})\leqslant\lambda$，$\forall\, i\in I$，因而

$$\mathrm{epi}(f)=\bigcap_{i\in I}\mathrm{epi}(f_i).$$

因为凸集 $\mathrm{epi}(f_i)$ 的交是凸集，所以 $f(\boldsymbol{x})$ 为凸函数.

性质5 设 $f: V\to\mathscr{R}\cup\{+\infty\}$ 为正常凸函数，$\boldsymbol{x}^{(i)}\in V$，$\lambda_i\geqslant 0$，$i=1，\cdots，m$，$\lambda_1+\cdots+\lambda_m=1$. 则

$$f\left(\sum_{i=1}^{m}\lambda_i\boldsymbol{x}^{(i)}\right)\leqslant\sum_{i=1}^{m}\lambda_i f(\boldsymbol{x}^{(i)}). \tag{3.3-4}$$

证明1 （用数学归纳法证，略）

证明2 （利用 $f(\boldsymbol{x})$ 的上图 $\mathrm{epi}(f)$）

因为 $f(\boldsymbol{x})$ 为凸函数，所以它的上图 $\mathrm{epi}(f)$ 为凸集.

$$(\boldsymbol{x}^{(i)},\ f(\boldsymbol{x}^{(i)}))\in\mathrm{epi}(f),\quad i=1，\cdots，m,$$

所以

$$\sum_{i=1}^{m}\lambda_i(\boldsymbol{x}^{(i)},\ f(\boldsymbol{x}^{(i)}))\in\mathrm{epi}(f),$$

即

$$\left(\sum_{i=1}^{m}\lambda_i\boldsymbol{x}^{(i)},\quad \sum_{i=1}^{m}\lambda_i f(\boldsymbol{x}^{(i)})\right)\in\mathrm{epi}(f),$$

于是得

$$f\left(\sum_{i=1}^{m}\lambda_i x^{(i)}\right)\leqslant\sum_{i=1}^{m}\lambda_i f(x^{(i)}).\qquad(3.3-5)$$

推论1 设$f:\mathscr{R}^n\to\mathscr{R}$为凸函数，$x^{(i)}\in\mathscr{R}^n$，$i=1,\cdots,m$，则

$$f\left(\frac{1}{m}(x^{(1)}+\cdots+x^{(m)})\right)\leqslant\frac{1}{m}\sum_{i=1}^{m}f(x^{(i)}).\qquad(3.3-6)$$

推论2 设$f:\mathscr{R}^n\to\mathscr{R}$为严格凸函数，$x^{(i)}\in\mathscr{R}^n$，$i=1,\cdots,m$. 若这$m$个点不完全相同，则下面的严格不等式成立

$$f\left(\frac{1}{m}(x^{(1)}+\cdots+x^{(m)})\right)<\frac{1}{m}\sum_{i=1}^{m}f(x^{(i)}).\qquad(3.3-7)$$

（证明从略）

性质6 设$f_i:V\to\mathscr{R}\cup\{+\infty\}$为正常凸函数，$i=1,2$，则

1° $\mathrm{dom}(f_1+f_2)=\mathrm{dom}(f_1)\cap\mathrm{dom}(f_2)$;

2° 若 $\mathrm{dom}(f_1)\cap\mathrm{dom}(f_2)\neq\Phi$，则$f(x)=f_1(x)+f_2(x)$为正常凸函数.

证 1° 因为$f_1(x)$，$f_2(x)$均不取$-\infty$，所以

$$f_1(x)+f_2(x)<+\infty\Longleftrightarrow f_1(x)<+\infty,\ f_2(x)<+\infty$$

因而

$$\mathrm{dom}(f_1+f_2)=\mathrm{dom}(f_1)\cap\mathrm{dom}(f_2).$$

2° 设x，$y\in V$，μ，$\nu\in\mathscr{R}$，$0<\lambda<1$，使

$$f(x)<\mu,\ f(y)<\nu.$$

令 μ_1，μ_2，ν_1，$\nu_2\in\mathscr{R}$，满足

$$f_i(x)<\mu_i,\ f_i(y)<\nu_i,\quad i=1,2,$$

$$\mu_1+\mu_2=\mu,\ \nu_1+\nu_2=\nu.$$

因为$f_i(x)$为凸函数，所以

$$f_i(\lambda x+(1-\lambda)y)<\lambda\mu_i+(1-\lambda)\nu_i,\quad i=1,2.$$

将上式对i相加得

$$f(\lambda \boldsymbol{x}+(1-\lambda)\boldsymbol{y})<\lambda\mu+(1-\lambda)\nu,$$

因而$f(\boldsymbol{x})$为凸函数.

又因为$f(\boldsymbol{x})$不取$-\infty$，而且$\mathrm{dom}(f)=\mathrm{dom}(f_1)\cap\mathrm{dom}(f_2)\neq\Phi$，所以$f(\boldsymbol{x})$为正常凸函数.

性质7 设$f:E\to\overline{\mathscr{R}}$为非正常凸函数，则

$$f(\boldsymbol{x})=-\infty,\ \forall \boldsymbol{x}\in \mathrm{ri}(\mathrm{dom}(f)).$$

证 若$\forall \boldsymbol{x}\in E\ f(\boldsymbol{x})=+\infty$,则命题显然成立. 否则,存在$\boldsymbol{a}\in E$，使$f(\boldsymbol{a})=-\infty$，$\boldsymbol{a}\in\mathrm{dom}(f)$. 再设$\boldsymbol{x}\in\mathrm{ri}(\mathrm{dom}(f))$，$\boldsymbol{x}\neq\boldsymbol{a}$，则必存在点$\boldsymbol{y}\in\mathrm{dom}(f)$与$\lambda\in(0,1)$使

$$\boldsymbol{x}=\lambda\boldsymbol{a}+(1-\lambda)\boldsymbol{y}.$$

因为$f(\boldsymbol{x})$为凸函数，而且$f(\boldsymbol{a})<\mu$，$\forall\mu\in\mathscr{R}$，所以对于满足$f(\boldsymbol{y})<\nu$的任意实数ν都有

$$f(\boldsymbol{x})=f(\lambda\boldsymbol{a}+(1-\lambda)\boldsymbol{y})<\lambda\mu+(1-\lambda)\nu$$

令$\mu\to-\infty$即得$f(\boldsymbol{x})=-\infty$.

性质8 设$f:\mathscr{R}\to\mathscr{R}\cup\{+\infty\}$为非减正常凸函数，$g:V\to\mathscr{R}$为凸函数，则复合函数$F=f(g):V\to\mathscr{R}\cup\{+\infty\}$为凸函数.

证 设$\boldsymbol{x}$，$\boldsymbol{y}\in V$，μ，$\nu\in\mathscr{R}$，$0<\lambda<1$，使

$$F(\boldsymbol{x})=f(g(\boldsymbol{x}))<\mu,\ F(\boldsymbol{y})=f(g(\boldsymbol{y}))<\nu.$$

因为$g(\boldsymbol{x})$是实值凸函数，所以

$$g(\lambda\boldsymbol{x}+(1-\lambda)\boldsymbol{y})\leqslant\lambda g(\boldsymbol{x})+(1-\lambda)g(\boldsymbol{y}).$$

因为$f(t)$非减，所以

$$f(g(\lambda\boldsymbol{x}+(1-\lambda)\boldsymbol{y}))\leqslant f(\lambda g(\boldsymbol{x})+(1-\lambda)g(\boldsymbol{y})).\quad(3.3\text{-}8)$$

再由$f(t)$的凸性得

$$\begin{aligned}&f(\lambda g(\boldsymbol{x})+(1-\lambda)g(\boldsymbol{y}))\\&<\lambda[f(g(\boldsymbol{x})+\varepsilon]++(1-\lambda)[f(g(\boldsymbol{y}))+\varepsilon]\\&=\lambda f(g(\boldsymbol{x}))+(1-\lambda)f(g(\boldsymbol{y}))+\varepsilon.\end{aligned}$$

因为其中$\varepsilon>0$可以任意小，所以

$$f(\lambda g(\boldsymbol{x})+(1-\lambda)g(\boldsymbol{y}))\leqslant\lambda f(g(\boldsymbol{x}))+(1-\lambda)f(g(\boldsymbol{y}))$$

$$<\lambda\mu+(1-\lambda)\nu. \qquad (3.3\text{-}9)$$

由式(3.3-8)与(3.3-9)得

$$f(g(\lambda \boldsymbol{x}+(1-\lambda)\boldsymbol{y}))<\lambda\mu+(1-\lambda)\nu,$$

即

$$F(\lambda \boldsymbol{x}+(1-\lambda)\boldsymbol{y})<\lambda\mu+(1-\lambda)\nu,$$

因而$F(\boldsymbol{x})=f(g(\boldsymbol{x}))$为凸函数.

(若$f(t)$为实值非减凸函数，则证明可以简化.)

*§3.4 凸函数的连续性

本节首先引入函数$f(\boldsymbol{x})$的下半连续性的概念，然后讨论函数的上图epi(f)为闭集的等价条件，最后研究凸函数的连续性.

3.4.1 函数的下半连续性

定义 3.4.1 设有函数$f:E\to\overline{\mathscr{R}}$，$\boldsymbol{a}\in E$. 若对于每个$k\in\mathscr{R}$：$k<f(\boldsymbol{a})$，都存在$\boldsymbol{a}$的邻域$U$，使$f(U)>k$，则称$f(\boldsymbol{x})$在点$\boldsymbol{a}$下半连续(lower semi-continuous).

若在集合$C\subset E$内每一点函数$f(\boldsymbol{x})$都是下半连续的，则称$f(\boldsymbol{x})$在C内下半连续.

定义 3.4.2 设有函数$f:E\to\overline{\mathscr{R}}$，$\boldsymbol{a}\in E$. 若对于每个$k\in\mathscr{R}$：$k>f(\boldsymbol{a})$，都存在$\boldsymbol{a}$的邻域$U$，使$f(U)<k$，则称$f(\boldsymbol{x})$在点$\boldsymbol{a}$上半续.

若在集合$C\subset E$内每一点$f(\boldsymbol{x})$都是上半连续的，则称$f(\boldsymbol{x})$在C内上半连续.

例 1 设f: $E\to\mathscr{R}$为连续函数，则$f(\boldsymbol{x})$既是下半连续函数，又是上半连续函数.

例 2 设$f:E\to\overline{\mathscr{R}}$，$a$，$b\in E$，且$f(\boldsymbol{a})=-\infty$，$f(\boldsymbol{b})=+\infty$，则$f(\boldsymbol{x})$在点$\boldsymbol{a}$是下半连续的，在点$\boldsymbol{b}$是上半连续的.

例 3 考虑§3.1例3中的函数$f(x)$. $f(x)$有两个间断点：

-1与1．$f(x)$在点$x=1$是下半连续的，不是上半连续的．$f(x)$在点$x=-1$是上半连续的，不是下半连续的．在其它点$f(x)$既是上半连续的，又是下半连续的．

定理 3.4.1 设$f: E\to\overline{\mathscr{R}}$，则下列条件等价：

1° $f(\boldsymbol{x})$在E上是下半连续的；

2° 对于每个$\lambda\in\mathscr{R}$，$A=\{\boldsymbol{x}\mid f(\boldsymbol{x})>\lambda\}$是开集；

3° 对于每个$\lambda\in\mathscr{R}$，$B=\{\boldsymbol{x}\mid f(\boldsymbol{x})\leqslant\lambda\}$是闭集；

4° $\mathrm{epi}(f)$是闭集．

证 "$1^\circ\Rightarrow2^\circ$" 设$f(\boldsymbol{x})$在$E$上是下半连续的．再设$\boldsymbol{a}\in A$，即$f(\boldsymbol{a})>\lambda$．根据定义3.4.1存在$\boldsymbol{a}$的邻域$U$，使$f(U)>\lambda$，因而$U\subset A$．所以$\boldsymbol{a}\in\mathrm{int}(A)$．由$\boldsymbol{a}$的任意性得$A$为开集．

"$2^\circ\Rightarrow3^\circ$" 设$A$是开集．因为$B=E\setminus A$，所以由$A$的开性得$B$的闭性．

"$3^\circ\Rightarrow4^\circ$" 设$B$是闭集．再设$(\boldsymbol{x}^{(k)},\ \lambda_k)\in\mathrm{epi}(f)$，$k=1, 2, \cdots$，并且

$$\lim_{k\to\infty}(\boldsymbol{x}^{(k)},\ \lambda_k)=(\hat{\boldsymbol{x}},\ \lambda).$$

用反证法证明$(\hat{\boldsymbol{x}},\ \lambda)\in\mathrm{epi}(f)$．倘若不然，即$f(\hat{\boldsymbol{x}})>\lambda$，于是存在$\varepsilon>0$，使$f(\hat{\boldsymbol{x}})>\lambda+\varepsilon$．

由假设知：存在K，使

$$f(\boldsymbol{x}^{(k)})\leqslant\lambda_k\leqslant\lambda+\varepsilon,\quad \forall k>K.$$

因而$\boldsymbol{x}^{(k)}\in\{\boldsymbol{x}\mid f(\boldsymbol{x})\leqslant\lambda+\varepsilon\}$，$\forall k>K$．再根据条件$3^\circ$得

$$\hat{\boldsymbol{x}}\in\{\boldsymbol{x}\mid f(\boldsymbol{x})\leqslant\lambda+\varepsilon\}$$

即$f(\hat{\boldsymbol{x}})\leqslant\lambda+\varepsilon$．这与$f(\hat{\boldsymbol{x}})>\lambda+\varepsilon$矛盾．所以$(\hat{\boldsymbol{x}},\ \lambda)\in\mathrm{epi}(f)$，即$\mathrm{epi}(f)$是闭集．

"$4^\circ\Rightarrow1^\circ$" 设$\mathrm{epi}(f)$是闭集．若$f(\boldsymbol{x})$在某一点$\boldsymbol{a}\in E$不是下半连续的，则对于某个$\beta<f(\boldsymbol{a})$，在$\boldsymbol{a}$的任一邻域$U_k$中都至少包含一点$\boldsymbol{x}^{(k)}$，使$f(\boldsymbol{x}^{(k)})\leqslant\beta$．因此可形成点列$\{\boldsymbol{x}^{(k)}\}\subset E$，$f(\boldsymbol{x}^{(k)})\leqslant$

β, $x^{(k)}\to a(k\to\infty)$. 于是$(x^{(k)},\ \beta)\in\text{epi}(f)$. 再根据epi$(f)$的闭性得$(a,\ \beta)\in\text{epi}(f)$, 即$f(a)\leqslant\beta$. 这与$\beta<f(a)$矛盾. 所以$f(x)$是下半连续的.

定理 3.4.2 设$f:E\to\overline{\mathscr{R}}$. 则下列条件等价.

1° $f(x)$在E上是上半连续的;

2° 对于每个$\lambda\in\mathscr{R}$, $A=\{x|\ f(x)<\lambda\}$是开集;

3° 对于每个$\lambda\in\mathscr{R}$, $B=\{x|\ f(x)\geqslant\lambda\}$是闭集;

4° 下图$\{(x,\ \lambda)\ |\lambda\leqslant f(x)\}$是闭集.

(证略)

若$f(x)$为连续性函数, 则不仅超平面$H=\{x\,|f(x)=\alpha\}$是闭的(定理1.5.4), 而且$\{x\,|f(x)\geqslant\alpha\}$与$\{x\,|f(x)\leqslant\alpha\}$都是闭的, $\{x\,|f(x)>\alpha\}$与$\{x\,|f(x)<\alpha\}$都是开的, 因此称前两个集合为闭半空间, 后两个集合为开半空间(定义1.5.2)是合理的.

定义 3.4.3 设$f:E\to\overline{\mathscr{R}}$, $a\in E$, U为a的邻域. 则称下式

$$\sup_U\{\inf f(x)\,|x\in U\setminus\{a\}\}\tag{3.4-1}$$

称为$f(x)$在点a的下极限, 记为 $\liminf\limits_{x\to a}f(x)$.

$(\{\inf f(x)\,|x\in U\setminus\{a\}\}$依赖于$U$)

定理 3.4.3 设$f:E\to\overline{\mathscr{R}}$, $a\in E$, 则$f(x)$在点a下半连续的充分必要条件是

$$\liminf_{x\to a} f(x)\geqslant f(a)\tag{3.4-2}$$

证 必要性. 设$f(x)$在点a下半连续, 而

$$\liminf_{x\to a} f(x)<f(a),$$

由此得

$$\liminf_{x\to a} f(x)+\frac{1}{2}(f(a)-\liminf_{x\to a} f(x))<f(a)$$

根据定义3.4.1存在a的邻域U, 使

$$f(U) > \liminf_{x \to a} f(x) + \frac{1}{2}(f(a) - \liminf_{x \to a} f(x)).$$

由上式及定义3.4.3得

$$\liminf_{x \to a} f(x) \geqslant \liminf_{x \to a} f(x) + \frac{1}{2}(f(a) - \liminf_{x \to a} f(x))$$

$$= \frac{1}{2}(f(a) + \liminf_{x \to a} f(x)).$$

由此得

$$\liminf_{x \to a} f(x) \leqslant f(a).$$

这与假设矛盾. 所以式(3.4-2)成立. 必要性得证.

充分性. 设条件式(3.4-2)成立, 而$f(x)$在点a不是下半连续的. 即对于某个k: $k < f(a)$, 在a的任一邻域U中至少存在一点x, 使$f(x) \leqslant k$, 因而

$$\{\inf f(x) \mid x \in U \setminus \{a\}\} \leqslant k.$$

于是 $\liminf_{x \to a} f(x) \leqslant k < f(a)$. 这与条件(3.4-2)矛盾. 所以$f(x)$在点$a$是下半连续的.

定义 3.4.4 设$A \subset V$, 则称

$$\delta_A(x) = \begin{cases} 0, & x \in A; \\ +\infty, & x \in V \setminus A \end{cases} \tag{3.4-3}$$

为A的**指示函数**(indicator function).

定理 3.4.4 设$A \subset V$, $\delta_A(x)$为A的指示函数, 则

1° $\delta_A(x)$为正常凸函数的充分必要条件是: A为非空凸集;

2° $\delta_A(x)$为下半连续函数的充分必要条件是: A为闭集.

证 1° 必要性. 若$\delta_A(x)$为正常凸函数, 则它的有效域$\mathrm{dom}(\delta_A)$(即A)为凸集, 而且不空.

充分性. 若A为非空凸集, 则定义在A上的零值函数$f(x) = 0$是凸函数, 因而$\delta_A(x)$是正常凸函数.

2° 必要性．因为$\delta_A(\boldsymbol{x})$是下半连续的，所以根据定理3.4.1 $A=\{\boldsymbol{x}\}|\delta_A(\boldsymbol{x})\leqslant 0\}$是闭集．

充分性．设A为闭集．对于$\lambda\in[0,+\infty)$，$\{\boldsymbol{x}|\delta_A(\boldsymbol{x})\leqslant\lambda\}=A$，所以$\{\boldsymbol{x}|\delta_A(\boldsymbol{x})\leqslant\lambda\}$是闭集．对于$\lambda\in(-\infty,0)$，$\{\boldsymbol{x}|\delta_A(\boldsymbol{x})\leqslant\lambda\}=\Phi$，空集$\Phi$也是闭集．根据定理3.4.1 $\delta_A(\boldsymbol{x})$是下半连续的．

定理 3.4.5 设$f:E\to\overline{\mathscr{R}}$是非正常凸函数，而且是下半连续的，则

$$f(\boldsymbol{x})=-\infty,\quad \forall\boldsymbol{x}\in\mathrm{dom}(f),$$

而且$\mathrm{dom}(f)$是闭集．

证 若$f(\boldsymbol{x})=+\infty$，$\forall\boldsymbol{x}\in E$，则$\mathrm{dom}(f)=\Phi$，命题显然成立．否则，存在点$\boldsymbol{a}\in\mathrm{dom}(f)$，$f(\boldsymbol{a})=-\infty$．假设存在点$\boldsymbol{b}\in\mathrm{dom}(f)$，$f(\boldsymbol{b})\in\mathscr{R}$．令$m$为通过$\boldsymbol{a}$与$\boldsymbol{b}$的直线．$f(\boldsymbol{x})$在直线$m$上的限制$f|m$是凸的（根据定理3.2.1），而且是非正常的．因为凸函数的有效域是凸的，所以

$$[\boldsymbol{a},\boldsymbol{b}]\subset\mathrm{dom}(f|m).$$

显然

$$(\boldsymbol{a},\boldsymbol{b})\subset\mathrm{int}(\mathrm{dom}(f|m)).$$

根据§3.3性质7，$f(\boldsymbol{x})=-\infty$，$\forall\boldsymbol{x}\in(a,b)$．这与$f(\boldsymbol{x})$在点$\boldsymbol{b}$的下半连续性矛盾，所以

$$f(\boldsymbol{x})=-\infty,\quad\forall\boldsymbol{x}\in\mathrm{dom}(f).$$

因为$f(\boldsymbol{x})$是下半连续的，所以根据定理3.4.1有效域$\mathrm{dom}(f)=\{\boldsymbol{x}|f(\boldsymbol{x})\leqslant 0\}$是闭集．

3.4.2 凸函数的连续性

定义 3.4.5 设$f:E\to\overline{\mathscr{R}}$，$\boldsymbol{a}\in E$．若存在点$\boldsymbol{a}$的一个邻域$U$与$M\in\mathscr{R}$，使$f(U)\leqslant M$，则称$f(\boldsymbol{x})$在点$\boldsymbol{a}$**局部上有界**．

定理 3.4.6 设$f:E\to\overline{\mathscr{R}}$是凸函数，$\boldsymbol{a}\in E$，$f(\boldsymbol{x})$在点$\boldsymbol{a}$局部上有界，并且$f(\boldsymbol{a})\in\mathscr{R}$，则

1° $f(\boldsymbol{x})$是正常凸函数；且$\mathrm{int}(\mathrm{dom}(f))\neq\Phi$；

2° 在int(dom(f))的每一点$f(\boldsymbol{x})$都是局部上有界的；

3° 在int(dom(f))的每一点$f(\boldsymbol{x})$都是连续的.

证 不妨设$\boldsymbol{a}=0$（否则，可令$\boldsymbol{t}=\boldsymbol{x}-\boldsymbol{a}$，将$f(\boldsymbol{x})$化为$\varphi(\boldsymbol{t})=f(\boldsymbol{t}+\boldsymbol{a})$讨论）. 因为$f(\boldsymbol{x})$在点0局部上有界，所以存在点0的邻域$U$与$M\in\mathscr{R}$，使$f(U)\leqslant M$. 因而$U\subset\text{dom}(f)$.

1° 因为$0\in U\subset\text{dom}(f)$，所以$0\in\text{int}(\text{dom}(f))$，因而int(dom($f$))$\neq\Phi$. 又因为$f(0)\neq-\infty$，所以$f(\boldsymbol{x})$是正常凸函数（根据§3.3性质7）.

2° 设$\boldsymbol{x}^{(0)}\in\text{int}(\text{dom}(f))$. 于是存在$\lambda>1$，使$\lambda\boldsymbol{x}^{(0)}\in\text{dom}(f)$. 因为$U$是点0的邻域，所以

$$W=\boldsymbol{x}^{(0)}+\left(1-\frac{1}{\lambda}\right)U$$

是$\boldsymbol{x}^{(0)}$的邻域. 以下证明在邻域W上$f(\boldsymbol{x})$上有界.

设$\boldsymbol{x}\in W$，则存在$\boldsymbol{u}\in U$，使

$$\begin{aligned}\boldsymbol{x}&=\boldsymbol{x}^{(0)}+\left(1-\frac{1}{\lambda}\right)\boldsymbol{u}\\&=\frac{1}{\lambda}(\lambda\boldsymbol{x}^{(0)})+\left(1-\frac{1}{\lambda}\right)\boldsymbol{u}\end{aligned}.$$

因为在dom(f)上$f(\boldsymbol{x})$是实值凸函数，所以

$$\begin{aligned}f(\boldsymbol{x})&=f\left(\frac{1}{\lambda}(\lambda\boldsymbol{x}^{(0)})+\left(1-\frac{1}{\lambda}\right)\boldsymbol{u}\right)\\&\leqslant\frac{1}{\lambda}f(\lambda\boldsymbol{x}^{(0)})+\left(1-\frac{1}{\lambda}\right)f(\boldsymbol{u}).\end{aligned}$$

因为$f(\boldsymbol{u})\leqslant M$，所以$f(\boldsymbol{x})\leqslant M_1$，此处

$$M_1=\frac{1}{\lambda}f\left(\lambda\boldsymbol{x}^{(0)}\right)+\left(1-\frac{1}{\lambda}\right)M.$$

由$\boldsymbol{x}$的任意性得$f(W)\leqslant M_1$. 因而$f(\boldsymbol{x})$在点$\boldsymbol{x}^{(0)}$局部上有界.

3° 先证明$f(\boldsymbol{x})$在点0是连续的．设$\varepsilon\in(0,1)$．因为U是点0的邻域，所以

$$D=\varepsilon(U\cap(-U))$$

也是点0的邻域．若$\boldsymbol{x}\in D$，则$\frac{1}{\varepsilon}\boldsymbol{x}\in U\cap(-U)$．因而

$$\frac{1}{\varepsilon}\boldsymbol{x}\in U,\quad -\frac{1}{\varepsilon}\boldsymbol{x}\in U.$$

于是

$$f\left(\frac{1}{\varepsilon}\boldsymbol{x}\right)\leqslant M,\quad f\left(-\frac{1}{\varepsilon}\boldsymbol{x}\right)\leqslant M.$$

根据$f(\boldsymbol{x})$的凸性可得

$$\begin{aligned}f(\boldsymbol{x})&=f\left(\varepsilon\left(\frac{1}{\varepsilon}\boldsymbol{x}\right)+(1-\varepsilon)0\right)\\&\leqslant\varepsilon f\left(\frac{1}{\varepsilon}\boldsymbol{x}\right)+(1-\varepsilon)f(0)\\&\leqslant\varepsilon M+(1-\varepsilon)f(0),\end{aligned}$$

因而

$$f(\boldsymbol{x})-f(0)\leqslant\varepsilon(M-f(0)).\tag{3.4-4}$$

根据$f(\boldsymbol{x})$的凸性可得

$$\begin{aligned}f(0)&=f\left(\frac{1}{1+\varepsilon}\boldsymbol{x}+\frac{\varepsilon}{1+\varepsilon}\left(-\frac{1}{\varepsilon}\boldsymbol{x}\right)\right)\\&\leqslant\frac{1}{1+\varepsilon}f(\boldsymbol{x})+\frac{\varepsilon}{1+\varepsilon}f\left(-\frac{1}{\varepsilon}\boldsymbol{x}\right)\\&\leqslant\frac{1}{1+\varepsilon}f(\boldsymbol{x})+\frac{\varepsilon}{1+\varepsilon}M,\end{aligned}$$

因而

$$f(\boldsymbol{x})-f(0)\geqslant-\varepsilon(M-f(0)).\tag{3.4-5}$$

由式(3.4-4)与(3.4-5)得

$$|f(\boldsymbol{x})-f(0)|\leqslant \varepsilon(M-f(0)).$$

因为 $\varepsilon>0$ 是任意的，所以$f(\boldsymbol{x})$在点 0 是连续的. 因为前面已设 $\boldsymbol{a}=0$，所以$f(\boldsymbol{x})$在点$\boldsymbol{a}$连续.

若$\boldsymbol{x}^0\in\operatorname{int}(\operatorname{dom}(f))$，$\boldsymbol{x}^0\neq 0$（即$\boldsymbol{x}^0\neq\boldsymbol{a}$），则根据 2°，$f(\boldsymbol{x})$在$\boldsymbol{x}^0$局部上有界，因而$f(\boldsymbol{x}^0)\in\mathscr{R}$. 根据前面的证明可得$f(\boldsymbol{x})$在点$\boldsymbol{x}^0$连续.

定义 3.4.6 设$A\subset E$（赋范线性空间），$f:E\to\overline{\mathscr{R}}$，$f(\boldsymbol{x})$在$A$上取实值，即$f(A)\subset\mathscr{R}$. 若存在正数$K$使不等式

$$|f(\boldsymbol{x})-f(\boldsymbol{y})|\leqslant K\|\boldsymbol{x}-\boldsymbol{y}\|,\quad\forall\boldsymbol{x},\ \boldsymbol{y}\in A,$$

恒成立，则称$f(\boldsymbol{x})$为关于A的Lipschitz算子.

定义 3.4.7 设$A\subset E$（赋范线性空间），$f:E\to\overline{\mathscr{R}}$，$f(A)\subset\mathscr{R}$. 若对于每个$\boldsymbol{a}\in A$，存在$\boldsymbol{a}$的一个邻域$U$使$f(\boldsymbol{x})$关于$U\cap A$是Lipschitz算子，则称$f(\boldsymbol{x})$为关于$A$的局部Lipschitz算子.

***定理 3.4.7** 设E为赋范线性空间，$f:E\to\mathscr{R}$为正常凸函数. 若$f(\boldsymbol{x})$在E的某一点$\boldsymbol{d}$局部上有界，则$f(\boldsymbol{x})$关于$\operatorname{int}(\operatorname{dom}(f))$是局部Lipschitz算子.

证 设$\boldsymbol{a}\in\operatorname{int}(\operatorname{dom}(f))$，则存在闭球 $\overline{B}(\boldsymbol{a};\ r_0)\subset\operatorname{int}(\operatorname{dom}(f))$. 根据定理3.4.6 $f(\boldsymbol{x})$在$\overline{B}(\boldsymbol{a};r_0)$上连续. 因而存在$m,M\in\mathscr{R}$，使

$$m\leqslant f(\boldsymbol{x})\leqslant M,\quad\forall\boldsymbol{x}\in\overline{B}(\boldsymbol{a};\ r_0).$$

设$0<r<r_0$，$\boldsymbol{x}$，$\boldsymbol{y}\in B(\boldsymbol{a};\ r)$，$\boldsymbol{x}\neq\boldsymbol{y}$，$\|\boldsymbol{x}-\boldsymbol{y}\|=\sigma$，再令

$$\boldsymbol{z}=\boldsymbol{y}+\frac{r_0-r}{\sigma}(\boldsymbol{y}-\boldsymbol{x}).\tag{3.4-6}$$

因为

$$\|\boldsymbol{z}-\boldsymbol{a}\|\leqslant\|\boldsymbol{y}-\boldsymbol{a}\|+r_0-r\leqslant r_0,$$

所以 $\boldsymbol{z}\in\overline{B}(\boldsymbol{a};\ r_0)$，如图3-3所示.

由式(3.4-6)得

$$\sigma \boldsymbol{z}=(\sigma+r_0-r)\boldsymbol{y}-(r_0-r)\boldsymbol{x}.$$

令$\lambda=\sigma/(\sigma+r_0-r)$，由上式可解得

$$\boldsymbol{y}=\lambda \boldsymbol{z}+(1-\lambda)\boldsymbol{x}.$$

根据$f(\boldsymbol{x})$的凸性以及$\boldsymbol{x}$，$\boldsymbol{y}$，$\boldsymbol{z}\in\overline{B}(\boldsymbol{a};\ r_0)$得

$$f(\boldsymbol{y})\leqslant\lambda f(\boldsymbol{z})+(1-\lambda)f(\boldsymbol{x}).$$

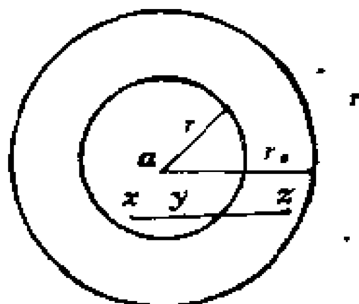


图 3-3

于是得

$$f(\boldsymbol{y})-f(\boldsymbol{x})\leqslant\lambda(f(\boldsymbol{z})-f(\boldsymbol{x}))\leqslant\lambda(M-m),$$

又因为

$$\lambda=\frac{\sigma}{\sigma+r_0-r}=\frac{\|\boldsymbol{x}-\boldsymbol{y}\|}{\sigma+r_0-r}\leqslant\frac{\|\boldsymbol{x}-\boldsymbol{y}\|}{r_0-r},$$

所以

$$f(\boldsymbol{y})-f(\boldsymbol{x})\leqslant\frac{M-m}{r_0-r}\|\boldsymbol{x}-\boldsymbol{y}\|.$$

在式(3.4-6)中交换$\boldsymbol{x}$，$\boldsymbol{y}$，可推得

$$f(\boldsymbol{x})-f(\boldsymbol{y})\leqslant\frac{M-m}{r_0-r}\|\boldsymbol{x}-\boldsymbol{y}\|.$$

令$K=(M-m)/(r_0-r)$，于是得

$$|f(\boldsymbol{y})-f(\boldsymbol{x})|\leqslant K\|\boldsymbol{y}-\boldsymbol{x}\|,\quad\forall\boldsymbol{x},\ \boldsymbol{y}\in B(\boldsymbol{a};\ r).$$

所以$f(\boldsymbol{x})$关于int(dom(f))是局部Lipschitz算子.

引理 设$A\subset\mathscr{R}^n$是开集，$\boldsymbol{x}^{(0)}\in A$. 则存在一个$n$维单纯形$S\subset A$，且$\boldsymbol{x}^{(0)}\in\text{int}(S)$.

证 由假设可知，存在$\varepsilon>0$使闭球$\overline{B}(\boldsymbol{x}^{(0)};\ \varepsilon)\subset A$. 设$\boldsymbol{e}^{(i)}$为$\mathscr{R}^n$中的单位向量，第$i$个分量为1，其余均为零. 再令

$$\boldsymbol{p}=-\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{e}^{(i)},$$

T为以$\varepsilon\boldsymbol{e}^{(1)}$，…，$\varepsilon\boldsymbol{e}^{(n)}$，$\varepsilon\boldsymbol{p}$为顶点的单纯形

$$T=\mathrm{co}(\varepsilon \boldsymbol{e}^{(1)},\ \cdots,\ \varepsilon \boldsymbol{e}^{(n)},\ \varepsilon \boldsymbol{p}).$$

显然$T\subset \overline{B}(0;\ \varepsilon)$. 因为

$$\frac{1}{2}\varepsilon \boldsymbol{p}+\sum_{i=1}^{n}\frac{1}{2n}\varepsilon \boldsymbol{e}^{(i)}=0,$$

所以$0\in \mathrm{int}(T)$（见第一章习题9）. 单纯形$S=T+\boldsymbol{x}^{(0)}$包含在闭球$\overline{B}(\boldsymbol{x}^{(0)};\ \varepsilon)$中，因而$S\subset A$，而且$\boldsymbol{x}^{(0)}\in \mathrm{int}(S)$(因为$0\in \mathrm{int}(T)$).

定理 3.4.8 设$f:\mathscr{R}^n\to\overline{\mathscr{R}}$为凸函数，则$f(\boldsymbol{x})$在$\mathrm{ri}(\mathrm{dom}(f))$上的限制是连续的.

证 若$f(\boldsymbol{x})$为非正常凸函数，则

$$f(\boldsymbol{x})=-\infty,\quad \forall\, \boldsymbol{x}\in \mathrm{ri}(\mathrm{dom}(f)).$$

因此$f(\boldsymbol{x})$在$\mathrm{ri}(\mathrm{dom}(f))$上的限制是连续的.

设$f(\boldsymbol{x})$为正常凸函数，$\boldsymbol{x}^{(0)}\in \mathrm{ri}(\mathrm{dom}(f))$, $\dim(\mathrm{ri}(\mathrm{dom}(f)))=k$. 根据引理，存在一个$k$维单纯形$S=\mathrm{co}(\boldsymbol{a}^{(0)},\ \boldsymbol{a}^{(1)},\ \cdots,\ \boldsymbol{a}^{(k)})$, $S\subset \mathrm{ri}(\mathrm{dom}(f))$，而且$\boldsymbol{x}^{(0)}\in \mathrm{ri}(S)$.

设$\boldsymbol{x}\in S$，即存在$\lambda_i\geqslant 0$，$\sum_{i=0}^{k}\lambda_i=1$，使

$$\boldsymbol{x}=\sum_{i=0}^{k}\lambda_i \boldsymbol{a}^{(i)}.$$

根据$f(\boldsymbol{x})$的凸性得

$$f(\boldsymbol{x})\leqslant \sum_{i=0}^{k}\lambda_i f(\boldsymbol{a}^{(i)}).$$

再令

$$M=\max\{f(\boldsymbol{a}^{(0)}),\ f(\boldsymbol{a}^{(1)}),\ \cdots,\ f(\boldsymbol{a}^{(k)})\},$$

则

$$f(\boldsymbol{x})\leqslant M,\quad \forall\, \boldsymbol{x}<S.$$

因而$f(\boldsymbol{x})$在$\mathrm{ri}(\mathrm{dom}(f))$上的限制（仍然是正常凸函数）在$\boldsymbol{x}^{(0)}$是局部上有界的. 根据定理3.4.6 $f(\boldsymbol{x})$在$\mathrm{ri}(\mathrm{dom}(f))$上的限制在$\mathrm{ri}(\mathrm{dom}(f))$的每一点都是连续的.

推论1 $\mathscr{R}^n$上的实凸函数$f(\boldsymbol{x})$都是连续函数.

（因为$\mathrm{ri}(\mathrm{dom}(f))=\mathscr{R}^n$）

推论2 设$f: \mathscr{R}^n \to \overline{\mathscr{R}}$为凸函数，$\boldsymbol{x}^{(0)} \in \mathrm{ri}(\mathrm{dom}(f))$，则$f(\boldsymbol{x})$在$\boldsymbol{x}^{(0)}$是下半连续的.

证 若$f(\boldsymbol{x})$为非正常凸函数，则$f(\boldsymbol{x}^{(0)})=-\infty$（见§3.3性质7）. 因而$f(\boldsymbol{x})$在$\boldsymbol{x}^{(0)}$是下半连续的.

若$f(\boldsymbol{x})$为正常凸函数，则$f(\boldsymbol{x})$在$\mathrm{ri}(\mathrm{dom}(f))$上的限制是连续的. 又因为$\boldsymbol{x}^{(0)} \in \mathrm{ri}(\mathrm{dom}(f))$，所以对任一$\mu \in \mathscr{R}$，$\mu<f(\boldsymbol{x}^{(0)})$，都存在$\boldsymbol{x}^{(0)}$的邻域$U$，使

$$U=(U \cap \mathrm{ri}(\mathrm{dom}(f))) \cup (U \setminus \mathrm{dom}(f)),$$

而且

$$f(\boldsymbol{x})>\mu, \quad \forall \boldsymbol{x} \in U \cap \mathrm{ri}(\mathrm{dom}(f)),$$

$$f(\boldsymbol{x})=+\infty, \quad \forall \boldsymbol{x} \in U \setminus \mathrm{dom}(f).$$

因而$\forall \boldsymbol{x} \in U$，都有$f(\boldsymbol{x})>\mu$. 所以$f(\boldsymbol{x})$在$\boldsymbol{x}^{(0)}$是下半连续的.

注 凸函数$f: \mathscr{R}^n \to \overline{\mathscr{R}}$在$\mathrm{ri}(\mathrm{dom}(f))$上不一定是连续的. 反例如下. 设$f: \mathscr{R}^2 \to \mathscr{R} \cup \{+\infty\}$

$$f(\boldsymbol{x})=\begin{cases} x_2^2, & x_1=0; \\ +\infty, & x_1 \neq 0. \end{cases}$$

则$\mathrm{ri}(\mathrm{dom}(f))=\{\boldsymbol{x} \mid x_1=0\}$. $\forall \boldsymbol{x} \in \mathrm{ri}(\mathrm{dom}(f))$，$f(\boldsymbol{x})$都不是连续的.

*§3.5 函数的下半连续包·凸包·下卷积

3.5.1 函数的下半连续包

定义 3.5.1 设$f, g: V \to \overline{\mathscr{R}}$. 若$g(\boldsymbol{x}) \leqslant f(\boldsymbol{x})$，$\forall \boldsymbol{x} \in V$（亦即

epi(g)$\supset$epi(f))，则称$g(x)$为$f(x)$的弱函数(minorant)。

定义 3.5.2 设$f:E\to\overline{\mathscr{R}}$. 则上图为$\overline{\text{epi}(f)}$的函数称为$f(x)$的**下半连续包**(lower semi-continuous hull)，记为$\overline{f}$. 由定义$\text{epi}(\overline{f})=\overline{\text{epi}(f)}$.

定理 3.5.1 设$f:E\to\overline{\mathscr{R}}$，$\overline{f}$为$f(x)$的下半连续包. 则$\overline{f}$是$f(x)$的最大下半连续弱函数.

证 因为epi($\overline{f}$)是闭集，所以$\overline{f}$是下半连续的(根据定理3.4.1).

因为epi($\overline{f}$)$\supset$epi(f)，所以$\overline{f}(x)\leqslant f(x)$. 因而$\overline{f}$是$f(x)$的弱函数.

设$g(x)$为$f(x)$的下半连续弱函数，则$g(x)\leqslant f(x)$，且epi(g)是闭的. 因而

$$\text{epi}(f)\subset\text{epi}(g),$$

于是

$$\text{epi}(\overline{f})=\overline{\text{epi}(f)}\subset\text{epi}(g).$$

由此得$\overline{f}(x)\geqslant g(x)$. 所以$\overline{f}$是$f(x)$的最大下半连续弱函数.

定理 3.5.2 设$f:E\to\overline{\mathscr{R}}$，$a\in E$，$\overline{f}$为$f(x)$的下半连续包，则

1° $f(x)$是下半连续的$\Longleftrightarrow f(x)=\overline{f}(x)$；

2° $\overline{f}(a)=\min\{f(a),\ \liminf\limits_{x\to a} f(x)\}$；

3° $f(x)$在点a下半连续$\Longleftrightarrow\overline{f}(a)=f(a)$.

证 1° 因为$\text{epi}(\overline{f})=\overline{\text{epi}(f)}$，所以根据定理3.4.1得

$$f(x)\text{下半连续}\Longleftrightarrow \text{eqi}(f)=\text{epi}(\overline{f})$$

$$\Longleftrightarrow f(x)=\overline{f}(x).$$

因而 $f(x)$下半连续$\Longleftrightarrow f(x)=\overline{f}(x)$.

2° 因为epi($\overline{f}$)是闭的，所以$\overline{f}(\boldsymbol{x})$是下半连续的，因而

$$\overline{f}(\boldsymbol{a})\leqslant\liminf_{\boldsymbol{x}\to\boldsymbol{a}}\overline{f}(\boldsymbol{x}) \tag{3.5-1}$$

（根据定理3.4.3）. 又因为$\overline{f}(\boldsymbol{x})\leqslant f(\boldsymbol{x})$，所以$\overline{f}(\boldsymbol{a})\leqslant f(\boldsymbol{a})$，且

$$\liminf_{\boldsymbol{x}\to\boldsymbol{a}}\overline{f}(\boldsymbol{x})\leqslant\liminf_{\boldsymbol{x}\to\boldsymbol{a}} f(\boldsymbol{x}). \tag{3.5-2}$$

由式(3.5-1)与(3.5-2)得

$$\overline{f}(\boldsymbol{a})\leqslant\liminf_{\boldsymbol{x}\to\boldsymbol{a}} f(\boldsymbol{x}).$$

令

$$\mu=\min\{f(\boldsymbol{a}),\ \liminf_{\boldsymbol{x}\to\boldsymbol{a}} f(\boldsymbol{x})\},$$

则

$$\overline{f}(\boldsymbol{a})\leqslant\mu.$$

假设$\overline{f}(\boldsymbol{a})<\mu$，则存在$\lambda\in\mathscr{R}$，使$\overline{f}(\boldsymbol{a})<\lambda<\mu$. 因而必存在点$\boldsymbol{a}$的一个邻域$U$，使

$$\inf\{f(\boldsymbol{x})\mid\boldsymbol{x}\in U\setminus\{\boldsymbol{a}\}\}>\lambda.$$

倘若不然，则 $\liminf_{\boldsymbol{x}\to\boldsymbol{a}} f(\boldsymbol{x})\leqslant\lambda$. 与$\lambda<\mu\leqslant\liminf_{\boldsymbol{x}\to\boldsymbol{a}} f(\boldsymbol{x})$矛盾.

另外还有$f(\boldsymbol{a})\geqslant\mu>\lambda$. 所以

$$f(\boldsymbol{x})>\lambda,\quad\forall\,\boldsymbol{x}\in U.$$

再令$I=(-\infty,\ \lambda)$，则$U\times I$是点$(\boldsymbol{a},\ \overline{f}(\boldsymbol{a}))$的一个邻域. 对于$U\times I$中的任一点$(\boldsymbol{x},\ \alpha)$均有$f(\boldsymbol{x})>\alpha$，因而不属于$f(\boldsymbol{x})$的上图

$$\mathrm{epi}(f)=\{(\boldsymbol{x},\ \alpha)\mid f(\boldsymbol{x})\leqslant\alpha\},$$

这与

$$(\boldsymbol{a},\ \overline{f}(\boldsymbol{a}))\in\mathrm{epi}(\overline{f})=\overline{\mathrm{epi}(f)}$$

矛盾，所以$\overline{f}(\boldsymbol{a})=\mu$，即

$$\overline{f}(\boldsymbol{a})=\min\{f(\boldsymbol{a}),\ \liminf_{\boldsymbol{x}\to\boldsymbol{a}} f(\boldsymbol{x})\}.$$

3° 因为

$$\overline{f}(\boldsymbol{a})=\min\{f(\boldsymbol{a}),\ \liminf_{\boldsymbol{x}\to\boldsymbol{a}} f(\boldsymbol{x})\},$$

所以

$$\overline{f}(a)=f(a)\Longleftrightarrow\varliminf_{x\to a} f(x)\geqslant f(a).$$

而

$$\varliminf_{x\to a} f(x)\geqslant f(a)$$

等价于$f(x)$在点a下半连续，所以

$f(x)$在点a下半连续$\Longleftrightarrow\overline{f}(a)=f(a)$.

例 1 设函数$f(x)$由式(3.1-7)给出. 求$f(x)$的下半连续包$\overline{f}(x)$.

解法一 在$x=-1$处$f(x)$不是下半连续的，在其余各点$f(x)$都是下半连续的. 所以只需考虑点$x=-1$. 设U为$x=-1$的任一邻域，则

$$\varliminf_{x\to -1} f(x)=\sup_{U}\{\inf f(x)\mid x\in U\setminus\{-1\}\}=1.$$

而$f(-1)=2$，$\min\{1,2\}=1$，因而根据定理3.5.2得$\overline{f}(-1)=1$. 于是得$f(x)$的下半连续包如下：

$$\overline{f}(x)=\begin{cases}x^2+1, & -\infty<x<-1;\\ x^2, & -1\leqslant x\leqslant 1;\\ x+1, & 1<x<+\infty.\end{cases}\qquad(3.5\text{-}3)$$

解法二 由图3-2知，epi(f)的闭包

$$\overline{\text{epi}(f)}=\text{epi}(f)\cup[c, d].$$

$f(x)$的下半连续包$\overline{f}(x)$，即以$\overline{\text{epi}(f)}$为上图的函数，就是式(3.5-3).

例2 设$f:V\to\overline{\mathscr{R}}$定义如下

$$f(x)=\inf\{\lambda\mid(x, \lambda)\in B\},$$

其中$B\subset V\times\mathscr{R}$为凸集. 试证明$f(x)$是上图包含$B$的最大凸函数.

证 因为已经证明$f(x)$是上图包含B的凸函数(见§3.1例4)，

所以只需证明：对于上图包含B的任一凸函数 $g(\boldsymbol{x})$，均有 $g(\boldsymbol{x})\leqslant f(\boldsymbol{x})$. 因为

$$g(\boldsymbol{x})=\inf\{\lambda\mid g(\boldsymbol{x})\leqslant\lambda\}=\inf\{\lambda\mid(\boldsymbol{x},\lambda)\in \mathrm{epi}(g)\},$$

$$f(\boldsymbol{x})=\inf\{\lambda\mid(\boldsymbol{x},\lambda)\in B\},$$

而且$\mathrm{epi}(g)\supset B$，所以$g(\boldsymbol{x})\leqslant f(\boldsymbol{x})$. 因而$f(\boldsymbol{x})$是上图包含 B 的最大凸函数.

定义 3.5.3 设$f(\boldsymbol{x})$为$\mathscr{R}^n$上的正常凸函数，则称$f(\boldsymbol{x})$的下半连续包$\overline{f}(\boldsymbol{x})$为$f(\boldsymbol{x})$的**闭包**(closure)，记为$\mathrm{cl}f(x)$.

定义 3.5.4 设$f(\boldsymbol{x})$为$\mathscr{R}^n$上的正常凸函数.若$\mathrm{cl}f(\boldsymbol{x})=f(\boldsymbol{x})$，则称$f(\boldsymbol{x})$为闭凸函数.

定理 3.5.3 设$f(\boldsymbol{x})$为$\mathscr{R}^n$上的正常凸函数，则$f(\boldsymbol{x})$为闭凸函数与$f(\boldsymbol{x})$下半连续等价.

证 设$f(\boldsymbol{x})$为闭凸函数. 根据定义得

$$\mathrm{epi}(f)=\mathrm{epi}(\mathrm{cl}f)=\overline{\mathrm{epi}(f)},$$

即$\mathrm{epi}(f)$是闭集. 由定理3.4.1知$f(\boldsymbol{x})$是下半连续的.

反之，设$f(\boldsymbol{x})$下半连续. 由定理3.4.1得$\mathrm{epi}(f)=\overline{\mathrm{epi}(f)}$，因而$\overline{f}=f$，即$\mathrm{cl}f=f$，所以$f(\boldsymbol{x})$为闭凸函数.

3.5.2 函数的凸包

定义 3.5.5 设$g:V\to\overline{\mathscr{R}}$为任一函数，则称

$$f(\boldsymbol{x})=\inf\{\lambda\mid(\boldsymbol{x},\lambda)\in \mathrm{co}(\mathrm{epi}(g))\} \qquad (3.5\text{-}4)$$

为**函数**$g(\boldsymbol{x})$的**凸包**(convex hull)，记为$\mathrm{co}(g)$.

定理 3.5.4 设$g:V\to\overline{\mathscr{R}}$为任一函数，$\mathrm{co}(g)$是$g(\boldsymbol{x})$的凸包，则$\mathrm{co}(g)$是$g(\boldsymbol{x})$的最大凸弱函数.

证 因为$\mathrm{co}(\mathrm{epi}(g))$是凸集，所以由例2知$\mathrm{co}(g)$是凸函数，而且

$$\mathrm{epi}(\mathrm{co}(g))\supset\mathrm{co}(\mathrm{epi}(g))\supset\mathrm{epi}(g),$$

所以$\mathrm{co}(g)(\boldsymbol{x})\leqslant g(\boldsymbol{x})$.

设$h(\boldsymbol{x})$为$g(\boldsymbol{x})$的任一凸弱函数，则$h(\boldsymbol{x})\leqslant g(\boldsymbol{x})$，因而$\mathrm{epi}(h)\supset\mathrm{epi}(g)$. 因为$\mathrm{epi}(h)$是凸集，所以

$$\mathrm{epi}(h)\supset \mathrm{co}(\mathrm{epi}(g)).$$

再根据co(g)的定义，以及

$$h(\boldsymbol{x})=\inf\{\lambda \mid (\boldsymbol{x}, \lambda)\in \mathrm{epi}(h)\}$$

得$h(\boldsymbol{x})\leqslant \mathrm{co}(g)(\boldsymbol{x})$，因而co($g$)是$g(\boldsymbol{x})$的最大凸弱函数

例3　设f_1，$f_2:\mathscr{R}\to\mathscr{R}$如下：

$$f_1(x)=\begin{cases}(x+1)^2, & x\in(-\infty, 0);\\ (x-1)^2, & x\in[0, +\infty).\end{cases}$$

$$f_2(x)=\begin{cases}e^x, & x\in(-\infty, 0);\\ (x-1)^2, & x\in[0, +\infty).\end{cases}$$

容易确定$f_1(x)$，$f_2(x)$的凸包如下：

$$\mathrm{co}(f_1)(x)=\begin{cases}(x+1)^2, & x\in(-\infty, -1);\\ 0, & x\in[-1, 1];\\ (x-1)^2, & x\in(1, +\infty).\end{cases}$$

$$\mathrm{co}(f_2)(x)=\begin{cases}0, & x\in(-\infty, 1);\\ (x-1)^2, & x\in[1, +\infty).\end{cases}$$

3.5.3 下卷积

定义 3.5.6　设$f, g:V\to\overline{\mathscr{R}}$均为凸函数．则称

$$\inf\{\lambda \mid (\boldsymbol{x}, \lambda)\in \mathrm{epi}(f)+\mathrm{epi}(g)\} \tag{3.5-5}$$

为$f(\boldsymbol{x})$与$g(\boldsymbol{x})$的**下卷积**(infimal convolution)，记为$(f\square g)(\boldsymbol{x})$，即

$$(f\square g)(\boldsymbol{x})=\inf\{\lambda \mid (\boldsymbol{x}, \lambda)\in \mathrm{epi}(f)+\mathrm{epi}(g)\}.$$

因为$f(\boldsymbol{x})$，$g(\boldsymbol{x})$均为凸函数，epi(f)，epi(g)均为凸集，epi(f)+epi(g)也是凸集，所以根据例2，下卷积$(f\square g)(\boldsymbol{x})$也是凸函数．容易证明

$$\mathrm{dom}(f\square g)=\mathrm{dom}(f)+\mathrm{dom}(g).$$

若$f(\boldsymbol{x})$，$g(\boldsymbol{x})$均为正常凸函数，则$f(\boldsymbol{x})$，$g(\boldsymbol{x})$的下卷积也可以有如下定义．

定义 3.5.7 设$f, g: V\to \mathscr{R}\cup\{+\infty\}$均为正常凸函数，则$f(\boldsymbol{x})$，$g(\boldsymbol{x})$的下卷积可定义如下：

$$(f\Box g)(\boldsymbol{x})=\inf\{f(\boldsymbol{x}^{(1)})+g(\boldsymbol{x}^{(2)})\mid \boldsymbol{x}^{(1)}+\boldsymbol{x}^{(2)}=\boldsymbol{x}\}.$$

可以证明，若f，$g: V\to \mathscr{R}\cup\{+\infty\}$均为正常凸函数，则以上两种定义等价.

上式可以改写成以下形式

$$(f\Box g)(\boldsymbol{x})=\inf_{\boldsymbol{y}\in V}\{f(\boldsymbol{x}-\boldsymbol{y})+g(\boldsymbol{y})\}.$$

例4 设$C\subset E$（赋范线性空间）为非空凸集，$\boldsymbol{x}\in E$，$f(\boldsymbol{x})=\|\boldsymbol{x}\|$，$\delta_C(\boldsymbol{x})$为$C$的指示函数，则

$$d(\boldsymbol{x}, C)=(f\Box\delta_C)(\boldsymbol{x}).$$

证

$$\begin{aligned} d(\boldsymbol{x}, C)&=\inf_{\boldsymbol{y}\in C}\|\boldsymbol{x}-\boldsymbol{y}\| \\ &=\inf_{\boldsymbol{y}\in E}\{\|\boldsymbol{x}-\boldsymbol{y}\|+\delta_C(\boldsymbol{y})\}=(f\Box\delta_C)(\boldsymbol{x}), \end{aligned}$$

所以

$$d(\boldsymbol{x}, C)=(f\Box\delta_C)(\boldsymbol{x}).$$

*§3.6 凸函数的可微性

本节首先讨论$\mathscr{R}$上的凸函数的左导数和右导数，然后讨论线性空间V中的凸函数的方向导数，最后简要地叙述函数$f(\boldsymbol{x})$的Fr'echet导数与Fr'echet微分.

3.6.1 左导数与右导数

定理 3.6.1 设I为$\mathscr{R}$中的一个区间，$f: I\to\mathscr{R}$为凸函数，$[a, b]\subset I$，则

$$1^\circ\quad f(x)\leqslant\frac{b-x}{b-a}f(a)+\frac{x-a}{b-a}f(b),\ \forall x\in[a, b];$$

$$2^\circ\quad \frac{f(x)-f(a)}{x-a}\leqslant\frac{f(b)-f(a)}{b-a}\leqslant\frac{f(b)-f(x)}{b-x},$$

$$\forall x\in(a,\ b).$$

（若$f(x)$为严格凸函数，则以上不等式均为严格不等式）

证 1° 因为$f(x)$为凸函数，所以

$$f(\lambda a+(1-\lambda)b)\leqslant\lambda f(a)+(1-\lambda)f(b),\quad \forall\lambda\in(0,\ 1).$$

令$x=\lambda a+(1-\lambda)b$，解出$\lambda=(b-x)/(b-a)$，于是$1-\lambda=(x-a)/(b-a)$. 将这些代入上式即得欲证.

2° 由1°得

$$f(x)\leqslant f(a)+\frac{x-a}{b-a}(f(b)-f(a)),$$

所以

$$\frac{f(x)-f(a)}{x-a}\leqslant\frac{f(b)-f(a)}{b-a}.$$

由1°还可得

$$f(x)\leqslant f(b)+\frac{b-x}{b-a}(f(a)-f(b)),$$

即

$$\frac{f(b)-f(a)}{b-a}\leqslant\frac{f(b)-f(x)}{b-x}.$$

定理 3.6.2 设$I\subset\mathscr{R}$为一个区间，$f:I\to\mathscr{R}$为凸函数，则

1° 在$\mathrm{int}(I)$中每一点c，$f(x)$的左导数$f'_-(c)$和右导数$f'_+(c)$都存在，而且

$$f'_-(c)\leqslant f'_+(c);\tag{3.6-1}$$

2° 在$\mathrm{int}(I)$中$f'_-(x)$与$f'_+(x)$均非减；

3° $\forall c\in\mathrm{int}(I)$，有

$$f(x)\geqslant f(c)+f'_-(c)(x-c),\quad \forall x\in I;\tag{3.6-2}$$

$$f(x)\geqslant f(c)+f'_+(c)(x-c),\quad \forall x\in I.\tag{3.6-3}$$

证 1° 首先证明在点$c\in\mathrm{int}(I)$，$f(x)$的左导数$f'_-(c)$和

右导数$f'_+(c)$存在.

因为$c\in \mathrm{int}(I)$，所以存在$[a,b]\subset I$，$a<c<b$. 对于每一点$x\in(a,c)$，$x<c<b$. 根据定理3.6.1得

$$\frac{f(x)-f(c)}{x-c}=\frac{f(c)-f(x)}{c-x}\leqslant\frac{f(b)-f(c)}{b-c}$$

而且当x_1，$x_2\in(a,c)$，$x_1<x_2$时有

$$\frac{f(x_1)-f(c)}{x_1-c}\leqslant\frac{f(x_2)-f(c)}{x_2-c}.$$

因而差商函数$(f(x)-f(c))/(x-c)$非减而且有上界，所以左导数存在，

$$f'_-(c)=\lim_{x\to c-}\frac{f(x)-f(c)}{x-c}.$$

同理可证右导数$f'_+(c)$也存在.

2° 根据定理1，若$a<c<d<b$，则对于充分小的正数ε有以下不等式

$$\frac{f(c)-f(c-\varepsilon)}{\varepsilon}\leqslant\frac{f(c+\varepsilon)-f(c)}{\varepsilon}\leqslant\frac{f(d)-f(d-\varepsilon)}{\varepsilon}$$

$$\leqslant\frac{f(d+\varepsilon)-f(d)}{\varepsilon}.$$

令$\varepsilon\to0^+$得

$$f'_-(c)\leqslant f'_+(c)\leqslant f'_-(d)\leqslant f'_+(d). \qquad (3.6\text{-}4)$$

这就证明了$f_-(x)$，$f'_+(x)$均非减，而且

$$f'_-(c)\leqslant f'_+(c).$$

3° 当$x=c$时，不等式显然成立.

若$x>c$，则对于充分小的正数ε有

$$\frac{f(c-\varepsilon)-f(c)}{-\varepsilon}\leqslant\frac{f(x)-f(c)}{x-c}.$$

令 $\varepsilon \to 0^+$ 得

$$f'_-(c) \leqslant \frac{f(x)-f(c)}{x-c}.$$

由此可得不等式

$$f(x) \geqslant f(c)+f'_-(c)(x-c).$$

若 $x<c$，则

$$\frac{f(x)-f(c)}{x-c} \leqslant f'_-(c).$$

由此也可得同样结果．这就证明了不等式（3.6-2）．

同理可证不等式（3.6-3）。

定理 3.6.3 设 $I \subset \mathscr{R}$ 是一个区间，$f: I \to \mathscr{R}$ 为凸函数，那么

1° 若 $[a,b] \subset \operatorname{int}(I)$，则 $f(x)$ 关于 $[a,b]$ 是Lipschitz 算子；

2° 在开区间 $\operatorname{int}(I)$ 内，$f'_-(x)$ 是左连续的，$f'_+(x)$ 是右连续的；

3° $f(x)$ 最多只能具有可数多个不可导的点．

证 1° 因为 $[a,b] \subset \operatorname{int}(I)$，所以存在 c，$d \in I$，使 $[a,b] \subset \operatorname{int}(c, d)$．根据定理3.6.2，$\forall x, y \in [a,b]$，$x<y$，均有

$$f(y) \geqslant f(x)+f'_+(x)(y-x),$$

$$f(x) \geqslant f(y)+f'_-(y)(x-y).$$

由以上不等式可得

$$f'_+(x) \leqslant \frac{f(y)-f(x)}{y-x} \leqslant f'_-(y).$$

因为 $f'_-(x)$，$f'_+(x)$ 均非减，所以由上式可得

$$f'_+(a) \leqslant \frac{f(y)-f(x)}{y-x} \leqslant f'_-(b).$$

令 $K=\max\{|f'_+(a)|, |f'_-(b)|\}$，则由上式可得

$$|f(y)-f(x)|\leqslant K|y-x|,$$

$\forall x,y\in[a,b]$，上式都成立．所以$f(x)$关于$[a,b]$是Lipschitz算子．

2° 由1°得$f(x)$在int(I)中是连续的，所以对于所有的$x,z,y\in\text{int}(I)$，$x<z<y$，均有

$$\frac{f(x)-f(y)}{x-y}=\lim_{z\to y-}\frac{f(x)-f(z)}{x-z}.$$

根据定理3.6.2得

$$\frac{f(x)-f(z)}{x-z}\leqslant f'_-(z),$$

因而

$$\frac{f(x)-f(y)}{x-y}\leqslant\lim_{z\to y-}f'_-(z).$$

令$x\to y^-$，得

$$f'_-(y)\leqslant\lim_{z\to y-}f'_-(z).$$

因为$f'_-(x)$非减，所以$f'_-(z)\leqslant f'_-(y)$，因而

$$\lim_{z\to y-}f'_-(z)\leqslant f'_-(y).$$

于是得

$$\lim_{z\to y-}f'_-(z)=f'_-(y).$$

这就证明了左导数$f'_-(x)$是左连续的．同理可证右导数$f'_+(x)$的右连续性．

3° 设x，z，$y\in\text{int}(I)$，$x<z<y$．根据不等式(3.6-4)得

$$f'_+(x)\leqslant f'_-(z)\leqslant f'_+(y).$$

若$f'_+(x)$在点z连续，则

$$\lim_{x\to z-}f'_+(x)=f'_+(z)=\lim_{y\to z+}f'_+(y).$$

根据以上二式得

$$f'_-(z)=f'_+(z).$$

因而$f(x)$在点z可导. 由此可知，若$f(x)$在点z不可导，则$f'_+(x)$在点z必不连续. 而$f'_+(x)$是非减的，所以不连续点最多有可数多个. 因而$f(x)$最多只能具有可数多个不可导的点.

定理 3.6.4 设$f:\mathscr{R}\to\mathscr{R}\cup\{+\infty\}$为正常凸函数，则在$f(x)$的有效域$\mathrm{dom}(f)$的每一点，左导数$f'_-(x)$与右导数$f'_+(x)$都存在（可以是$+\infty$或$-\infty$），而且$f'_-(x)\leqslant f'_+(x)$.

证 $\mathscr{R}$上的正常凸函数$f(x)$的有效域$\mathrm{dom}(f)$是凸集，就是一个区间. 若$\mathrm{dom}(f)$是一个开区间，则由定理3.6.2可得要证的结论. 所以不妨设

$$\mathrm{dom}(f)=[a,\ b],$$

而且只需考虑在两个端点a与b的情形.

因为差商

$$\frac{f(x)-f(b)}{x-b},\ x\in(a,\ b)$$

非减，所以当$x\to b^-$时，它的极限$f'_-(b)$存在（或为$+\infty$).

当$x>b$时，$f(x)=+\infty$，因而

$$\frac{f(x)-f(b)}{x-b}=+\infty.$$

令$x\to b^+$，得$f'_+(b)=+\infty$，所以$f'_-(b)\leqslant f'_+(b)$.

同理，$f'_+(a)$存在（或为$-\infty$），$f'_-(a)=-\infty$，因而$f'_-(a)\leqslant f'_+(a)$.

3.6.2 方向导数

定义 3.6.1 设$f:V\to\overline{\mathscr{R}}$，$\boldsymbol{x}^{(0)}$，$\boldsymbol{x}\in V$，$f(\boldsymbol{x}^{(0)})$为有限值. 若极限

$$\lim_{\lambda\to 0^+}\frac{f(\boldsymbol{x}^{(0)}+\lambda\boldsymbol{x})-f(\boldsymbol{x}^{(0)})}{\lambda} \tag{3.6-5}$$

存在，则称此极限为函数$f(\boldsymbol{x})$在点$\boldsymbol{x}^{(0)}$沿方向$\boldsymbol{x}$的（单边）方向导数，记为$f'(\boldsymbol{x}^{(0)};\ \boldsymbol{x})$*.（极限（3.6-5）的值可以是$+\infty$或$-\infty$）

定理 3.6.5 设$f(\boldsymbol{x})$为V上的正常凸函数，$f(\boldsymbol{x}^{(0)})$为有限值，则

1° $\forall \boldsymbol{x}\in V$，$f'(\boldsymbol{x}^{(0)};\ \boldsymbol{x})$存在；

2° $f'(\boldsymbol{x}^{(0)};\ \boldsymbol{x})$为$V$上的正齐次凸函数.

证 1° 设$\boldsymbol{x}\in V$，$\varphi(t)=f(\boldsymbol{x}^{(0)}+t\boldsymbol{x})$，$t\in\mathscr{R}$. 因为$f(\boldsymbol{x})$为正常凸函数，所以单元函数$\varphi(t)$也是正常凸函数，且$\varphi(0)$为有限值，$0\in\mathrm{dom}(\varphi)$. 由定理3.6.4得$\varphi'_+(0)$存在. 而

$$f'(\boldsymbol{x}^{(0)};\ \boldsymbol{x})=\lim_{t\to 0+}\frac{f(\boldsymbol{x}^{(0)}+t\boldsymbol{x})-f(\boldsymbol{x}^{(0)})}{t}$$

$$=\lim_{t\to 0+}\frac{\varphi(t)-\varphi(0)}{t}=\varphi'_+(0),$$

所以$f'(\boldsymbol{x}^{(0)};\ \boldsymbol{x})$存在.

2° 由定义3.6.5知：当$\alpha>0$时有

$$f'(\boldsymbol{x}^{(0)};\ \alpha\boldsymbol{x})=\lim_{\lambda\to 0+}\frac{f(\boldsymbol{x}^{(0)}+\lambda\alpha\boldsymbol{x})-f(\boldsymbol{x}^{(0)})}{\lambda}$$

$$=\alpha\lim_{\lambda\alpha\to 0+}\frac{f(\boldsymbol{x}^{(0)}+\lambda\alpha\boldsymbol{x})-f(\boldsymbol{x}^{(0)})}{\lambda\alpha}$$

$$=\alpha f'(\boldsymbol{x}^{(0)};\ \boldsymbol{x}).$$

因此$f'(\boldsymbol{x}^{(0)};\ \boldsymbol{x})$是正齐次的.

由$f(\boldsymbol{x})$的凸性知，$\forall\alpha\in(0,\ 1)$，$\forall\boldsymbol{y},\ \boldsymbol{z}\in V$均有

$$f(\boldsymbol{x}^{(0)}+\lambda(\alpha\boldsymbol{y}+(1-\alpha)\boldsymbol{z}))-f(\boldsymbol{x}^{(0)})$$
$$=f(\alpha(\boldsymbol{x}^{(0)}+\lambda\boldsymbol{y})+(1-\alpha)(\boldsymbol{x}^{(0)}+\lambda\boldsymbol{z}))-f(\boldsymbol{x}^{(0)})$$
$$\leqslant\alpha f(\boldsymbol{x}^{(0)}+\lambda\boldsymbol{y})+(1-\alpha)f(\boldsymbol{x}^{(0)}+\lambda\boldsymbol{z})-f(\boldsymbol{x}^{(0)})$$

* 也可以将极限(3.6-5)记为$f'_+(\boldsymbol{x}^{(0)};\boldsymbol{x})$。

$$=\alpha[f(\boldsymbol{x}^{(0)}+\lambda\boldsymbol{y})-f(\boldsymbol{x}^{(0)})]+(1-\alpha)[f(\boldsymbol{x}^{(0)}+\lambda\boldsymbol{z})-f(\boldsymbol{x}^{(0)})].$$

因而

$$\begin{aligned}&f(\boldsymbol{x}^{(0)}+\lambda(\alpha\boldsymbol{y}+(1-\alpha)\boldsymbol{z}))-f(\boldsymbol{x}^{(0)})\\ \leqslant&\alpha[f(\boldsymbol{x}^{(0)}+\lambda\boldsymbol{y})-f(\boldsymbol{x}^{(0)})]\\ &+(1-\alpha)[f(\boldsymbol{x}^{(0)}+\lambda\boldsymbol{z})-f(\boldsymbol{x}^{(0)})].\end{aligned}$$

两端除以λ，然后令$\lambda\to0^+$得

$$f'(\boldsymbol{x}^{(0)};\ \alpha\boldsymbol{y}+(1-\alpha)\boldsymbol{z})\leqslant\alpha f'(\boldsymbol{x}^{(0)};\boldsymbol{y})+(1-\alpha)f'(\boldsymbol{x}^{(0)};\ \boldsymbol{z}).$$

所以$f'(\boldsymbol{x}^{(0)};\ \boldsymbol{x})$为$V$上的正齐次凸函数.

3.6.3 Fr'echet微分

Fr'echet微分就是数学分析中多元函数全微分概念的推广. 其定义如下.

定义 3.6.2 设E为赋范线性空间，E'为E的对偶. $f:E\to\overline{\mathscr{R}}$，$\boldsymbol{x}^{(0)}\in E$，$f(\boldsymbol{x}^{(0)})\in\mathscr{R}$. 若存在$\boldsymbol{u}\in E'$，使

$$f(\boldsymbol{x})=f(\boldsymbol{x}^{(0)})+(\boldsymbol{x}-\boldsymbol{x}^{(0)}|\boldsymbol{u})+o(\|\boldsymbol{x}-\boldsymbol{x}^{(0)}\|),\quad(3.6\text{-}6)$$

亦即

$$\lim_{x\to x^{(0)}}\frac{f(\boldsymbol{x})-f(\boldsymbol{x}^{(0)})-(\boldsymbol{x}-\boldsymbol{x}^{(0)}|\boldsymbol{u})}{\|\boldsymbol{x}-\boldsymbol{x}^{(0)}\|}=0,\quad(3.6\text{-}7)$$

则称$\boldsymbol{u}$为$f(\boldsymbol{x})$在点$\boldsymbol{x}^{(0)}$的Fr'echet**导数**，记为$\nabla f(x^{(0)})$（也可称为$f(\boldsymbol{x})$在点$\boldsymbol{x}^{(0)}$的梯度）. 而$(\boldsymbol{x}-\boldsymbol{x}^{(0)}|\boldsymbol{u})$称为$f(\boldsymbol{x})$在点$\boldsymbol{x}^{(0)}$的Fr'echet**微分**，记为$df(\boldsymbol{x}^{(0)})$.

若函数$f(\boldsymbol{x})$在点$\boldsymbol{x}^{(0)}$的Fr'echet导数$\nabla f(\boldsymbol{x}^{(0)})$存在，则称$f(\boldsymbol{x})$在点$\boldsymbol{x}^{(0)}$ Fr'echet**可微**（简称可微）. 若$f(\boldsymbol{x})$在$C\subset E$上每点都可微，则称$f(\boldsymbol{x})$在C上可微.

若$f(\boldsymbol{x})$为$\mathscr{R}^n$中的实函数，$\boldsymbol{x}^{(0)}\in\mathscr{R}^n$，则$f(\boldsymbol{x})$在点$\boldsymbol{x}^{(0)}$的Fr'echet导数即数学分析中讨论的$f(\boldsymbol{x})$在点$\boldsymbol{x}^{(0)}$的梯度

$$\nabla f(\boldsymbol{x}^{(0)})=\left(\frac{\partial f}{\partial x_1},\ \cdots,\ \frac{\partial f}{\partial x_n}\right)^T.$$

而$f(\boldsymbol{x})$在点$\boldsymbol{x}^{(0)}$的Fr'echet微分即数学分析中的全微分

$$df(x^{(0)}) = \frac{\partial f}{\partial x_1}(x_1 - x_1^{(0)}) + \cdots + \frac{\partial f}{\partial x_n}(x_n - x_n^{(0)}).$$

定理 3.6.6 若$f: E \to \overline{\mathscr{R}}$在点$x^{(0)} \in E$可微，则$f(x)$在点$x^{(0)}$沿任一向量$y \in E$的方向导数存在，且

$$f'(x^{(0)};\ y) = (y \mid \nabla f(x^{(0)})). \tag{3.6-8}$$

证 令$x = x^{(0)} + ty$，则

$$\lim_{t \to 0+} \frac{f(x^{(0)} + ty) - f(x^{(0)})}{t}$$

$$= \lim_{x \to x^{(0)}} \frac{f(x) - f(x^{(0)}) - (x - x^{(0)} \mid \nabla f(x^{(0)}))}{\| x - x^{(0)} \|} \| y \|$$

$$+ (y \mid \nabla f(x^{(0)})) = (y \mid \nabla f(x^{(0)}))\text{（由式(3.6-7)）}.$$

于是得：$$f'(x^{(0)};\ y) = (y \mid \nabla f(x^{(0)})).$$

*§3.7 次梯度与次微分

本节研究有关凸函数$f(x)$的两个重要概念：次梯度与次微分。

定义 3.7.1 设$f: E \to \overline{\mathscr{R}}$为凸函数，$x^{(0)} \in E$，$f(x^{(0)}) \in \mathscr{R}$．若$u \in E'$使下面的不等式成立

$$f(x) \geqslant f(x^{(0)}) + (x - x^{(0)} \mid u),\quad \forall x \in E$$

则称u为$f(x)$在点$x^{(0)}$的一个**次梯度**（subgradient）.

若$E = \mathscr{R}^n$，则$f(x)$在点$x^{(0)}$的次梯度是$\mathscr{R}^n$中的一个向量.

定义 3.7.2 $f(x)$在点$x^{(0)}$的所有次梯度的集合，称为$f(x)$在点$x^{(0)}$的**次微分**（subdifferential），记为$\partial f(x^{(0)})$.

若$\partial f(x^{(0)}) \neq \Phi$，则称$f(x)$在点$x^{(0)}$是次可微的．若在$C \subset E$内每一点$x$，次微分$\partial f(x)$均非空，则称$f(x)$在$C$内是次可微的.

容易验证，次微分$\partial f(x^{(0)})$是E'中的一个闭凸集.

定理 3.7.1 设$f:(a,b)\to\mathscr{R}$为单元凸函数，则在(a,b)内$f(x)$是次可微的．若$c\in(a,b)$，则

$$\partial f(c)=[f'_-(c),\ f'_+(c)]. \tag{3.7-1}$$

证 根据定理3.6.2凸函数$f(x)$在(a,b)内每一点c的左导数$f'_-(c)$与右导数$f'_+(c)$都存在，而且

$$f(x)\geqslant f(c)+f'_-(c)(x-c),\quad \forall x\in(a,\ b)$$

$$f(x)\geqslant f(c)+f'_+(c)(x-c),\quad \forall x\in(a,\ b)$$

所以$f(x)$在点c是次可微的．

因为$f'_-(c)$，$f'_+(c)\in\partial f(c)$，而且$\partial f(c)$是凸集，所以

$$[f'_-(c),\ f'_+(c)]\subset\partial f(c). \tag{3.7-2}$$

设$t\in\partial f(c)$，即

$$f(x)\geqslant f(c)+t(x-c),\quad \forall x\in(a,\ b).$$

当$x>c$时，由上式可得

$$\frac{f(x)-f(c)}{x-c}\geqslant t\ .$$

令$x\to c^+$得$f'_+(c)\geqslant t$．同理可得$f'_-(c)\leqslant t$．

因而$t\in[f'_-(c),\ f'_+(c)]$．于是

$$[f'_-(c),\ f'_+(c)]\supset\partial f(c). \tag{3.7-3}$$

结合式（3.7-2）与式（3.7-3）得

$$\partial f(c)=[f'_-(c),\ f'_+(c)].$$

推论 若单元凸函数$f(x)$在点c可导，则

$$\partial f(c)=\{f'(c)\}.$$

例1 设单元凸函数如下：

$$f(x)=\begin{cases}e^{-x}, & x<0\ ;\\ e^{x}, & x\geqslant 0\ .\end{cases}$$

试研究$f(x)$的次可微性，并求次微分$\partial f(x)$．

解 根据定理3.7.1 $f(x)$在$(-\infty,+\infty)$内是次可微的．

当 $x\neq 0$ 时 $f(x)$ 可导，所以

$$\partial f(x)=\{f'(x)\}.$$

当 $x=0$ 时有 $f'_-(0)=-1$，$f'_+(0)=1$，因而

$$\partial f(0)=[-1,1].$$

（见图3-4，图中 $\boldsymbol{a}=(0,1)^T$）

定义 3.7.3 设 $f:E\to\overline{\mathscr{R}}$ 为凹函数，$\boldsymbol{x}^{(0)}\in E$，$f(\boldsymbol{x}^{(0)})\in\mathscr{R}$。若 $\boldsymbol{u}\in E'$ 使下面的不等式成立

$$f(\boldsymbol{x})\leqslant f(\boldsymbol{x}^{(0)})+(\boldsymbol{x}-\boldsymbol{x}^{(0)}|\boldsymbol{u}),\quad \forall\boldsymbol{x}\in E,$$

则称 $\boldsymbol{u}$ 为 $f(\boldsymbol{x})$ 在点 $\boldsymbol{x}^{(0)}$ 的一次梯度.

定理 3.7.2 设 $f:E\to\mathscr{R}$ 为凸函数，$\boldsymbol{x}^{(0)}\in E$，$f(\boldsymbol{x}^{(0)})\in\mathscr{R}$. 则 $\boldsymbol{u}\in E'$ 为 $f(\boldsymbol{x})$ 在点 $\boldsymbol{x}^{(0)}$ 的次梯度的充分必要条件是

$$f'(\boldsymbol{x}^{(0)};\ \boldsymbol{x})\geqslant(\boldsymbol{x}|\boldsymbol{u}),\quad\forall\boldsymbol{x}\in E.\tag{3.7-4}$$

证 必要性. 设 $\boldsymbol{u}\in\partial f(\boldsymbol{x}^{(0)})$，$\boldsymbol{x}\in E$，$t>0$. 根据定义3.7.1得

$$f(\boldsymbol{x}^{(0)}+t\boldsymbol{x})\geqslant f(\boldsymbol{x}^{(0)})+t(\boldsymbol{x}|\boldsymbol{u}),$$

因为 $t>0$，所以上式可化为

$$\frac{f(\boldsymbol{x}^{(0)}+t\boldsymbol{x})-f(\boldsymbol{x}^{(0)})}{t}\geqslant(\boldsymbol{x}|\boldsymbol{u}).$$

令 $t\to0^+$ 得

$$f'(\boldsymbol{x}^{(0)};\ \boldsymbol{x})\geqslant(\boldsymbol{x}|\boldsymbol{u}).$$

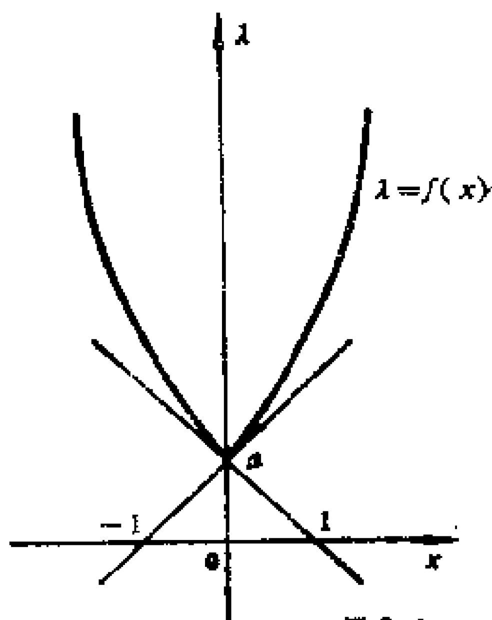


图 3-4

充分性. 设条件（3.7-4）成立，而 $\boldsymbol{u}\,\overline{\in}\,\partial f(\boldsymbol{x}^{(0)})$. 则必存在 $\boldsymbol{x}^{(1)}\in E$ 使

$$f(\boldsymbol{x}^{(1)})<f(\boldsymbol{x}^{(0)})+(\boldsymbol{x}-\boldsymbol{x}^{(0)}|\boldsymbol{u}).$$

取向量 $\boldsymbol{d}=\boldsymbol{x}^{(1)}-\boldsymbol{x}^{(0)}$，$0<t<1$. 根据 $f(\boldsymbol{x})$ 的凸性得

$$\frac{f(\boldsymbol{x}^{(0)}+t\boldsymbol{d})-f(\boldsymbol{x}^{(0)})}{t}$$

$$=\frac{f(t\boldsymbol{x}^{(1)}+(1-t)\boldsymbol{x}^{(0)})-f(\boldsymbol{x}^{(0)})}{t}$$

$$\leqslant\frac{tf(\boldsymbol{x}^{(1)})+(1-t)f(\boldsymbol{x}^{(0)})-f(\boldsymbol{x}^{(0)})}{t}$$

$$=f(\boldsymbol{x}^{(1)})-f(\boldsymbol{x}^{(0)})$$

$$<(\boldsymbol{x}^{(1)}-\boldsymbol{x}^{(0)}|\boldsymbol{u})=(\boldsymbol{d}|\boldsymbol{u}).$$

令 $t\to0^+$ 得

$$f'(\boldsymbol{x}^{(0)};\ \boldsymbol{d})\leqslant f(\boldsymbol{x}^{(1)})-f(\boldsymbol{x}^{(0)})<(\boldsymbol{d}|\boldsymbol{u}).$$

上式与条件（3.7-4）矛盾．所以 $\boldsymbol{u}\in\partial f(\boldsymbol{x}^{(0)})$．充分性得证．

定理 3.7.3 设 $f:E\to\mathscr{R}$ 为凸函数．$f(\boldsymbol{x})$ 在点 $\boldsymbol{x}^{(0)}$ 可微，$\nabla f(\boldsymbol{x}^{(0)})$ 为 $f(\boldsymbol{x})$ 在 $\boldsymbol{x}^{(0)}$ 的Fr'echet导数．则 $\nabla f(\boldsymbol{x}^{(0)})$ 是 $f(\boldsymbol{x})$ 在点 $\boldsymbol{x}^{(0)}$ 的唯一次梯度，即

$$\partial f(\boldsymbol{x}^{(0)})=\{\nabla f(\boldsymbol{x}^{(0)})\}. \qquad (3.7\text{-}5)$$

证 根据定理3.6.6，$\forall\boldsymbol{x}\in E$ 均有

$$f'(\boldsymbol{x}^{(0)};\ \boldsymbol{x})=(\boldsymbol{x}|\nabla f(\boldsymbol{x}^{(0)})). \qquad (3.7\text{-}6)$$

由定理3.7.2知，$\nabla f(\boldsymbol{x}^{(0)})\in\partial f(\boldsymbol{x}^{(0)})$．若 $\boldsymbol{u}\in\partial f(\boldsymbol{x}^{(0)})$，则由定理3.7.2，$\forall\boldsymbol{x}\in E$ 以下不等式

$$f'(\boldsymbol{x}^{(0)};\ \boldsymbol{x})\geqslant(\boldsymbol{x}|\boldsymbol{u}) \qquad (3.7\text{-}7)$$

恒成立．由式（3.7-6）与（3.7-7）得

$$(\boldsymbol{x}|\nabla f(\boldsymbol{x}^{(0)}))\geqslant(\boldsymbol{x}|\boldsymbol{u}),\qquad\forall\boldsymbol{x}\in E.$$

于是

$$(\boldsymbol{x}|(\nabla f(\boldsymbol{x}^{(0)})-\boldsymbol{u}))\geqslant0,\qquad\forall\boldsymbol{x}\in E.$$

若 $\boldsymbol{x}\in E$ 使上式成立（$\boldsymbol{x}\neq0$），则用 $-\boldsymbol{x}$ 代 $\boldsymbol{x}$ 上式也成立．所以 $\nabla f(\boldsymbol{x}^{(0)})-\boldsymbol{u}$ 为0，即 $\boldsymbol{u}=\nabla f(\boldsymbol{x}^{(0)})$．于是

$$\partial f(\boldsymbol{x}^{(0)})=\{\nabla f(\boldsymbol{x}^{(0)})\}.$$

定理 3.7.4 设 $f:E\to\mathscr{R}\cup\{+\infty\}$ 为正常凸函数，$\boldsymbol{x}^{(0)}\in\mathrm{dom}(f)$．则 $f(\boldsymbol{x})$ 在点 $\boldsymbol{x}^{(0)}$ 次可微的充分必要条件是：在点 $(\boldsymbol{x}^{(0)},\ f(\boldsymbol{x}^{(0)}))$ 对epi(f)的非垂直的闭的支撑超平面存在．

证 必要性. 设$f(\boldsymbol{x})$在点$\boldsymbol{x}^{(0)}$次可微，$\boldsymbol{u}\in\partial f(\boldsymbol{x}^{(0)})$. 令$\beta=(\boldsymbol{x}^{(0)}|\boldsymbol{u})-f(\boldsymbol{x}^{(0)})$，则$H$：

$$F(\boldsymbol{x},\lambda)=(\boldsymbol{x}|\boldsymbol{u})-\lambda=\beta \tag{3.7-8}$$

是一张非垂直的闭的超平面（参看式（1.5-1）），且$(\boldsymbol{x}^{(0)},f(\boldsymbol{x}^{(0)}))\in H$. 因为$\boldsymbol{u}\in\partial f(\boldsymbol{x}^{(0)})$，所以

$$f(\boldsymbol{x})\geqslant f(\boldsymbol{x}^{(0)})+(\boldsymbol{x}-\boldsymbol{x}^{(0)}|\boldsymbol{u}),\quad\forall\boldsymbol{x}\in E.$$

对于上图epi(f)中的点（$\boldsymbol{x}$，λ）均有

$$\begin{aligned}F(\boldsymbol{x},\lambda)&=(\boldsymbol{x}|\boldsymbol{u})-\lambda\leqslant(\boldsymbol{x}|\boldsymbol{u})-f(\boldsymbol{x})\\&\leqslant(\boldsymbol{x}|\boldsymbol{u})-f(\boldsymbol{x}^{(0)})-(\boldsymbol{x}-\boldsymbol{x}^{(0)}|\boldsymbol{u})\\&=(\boldsymbol{x}^{(0)}|\boldsymbol{u})-f(\boldsymbol{x}^{(0)})=\beta,\end{aligned}$$

因而$F(\boldsymbol{x},\lambda)\leqslant\beta$，所以$H$是在点$(\boldsymbol{x}^{(0)},f(\boldsymbol{x}^{(0)}))$对epi($f$)的非垂直的闭的支撑超平面.

充分性. 设H：

$$F(\boldsymbol{x},\lambda)=(\boldsymbol{x}|\boldsymbol{u})+\alpha\lambda=\beta,\quad\alpha\neq0$$

是在点$(\boldsymbol{x}^{(0)},f(\boldsymbol{x}^{(0)}))$对epi($f$)的一张非垂直的闭的支撑超平面. 不妨设$\forall(\boldsymbol{x},\lambda)\in$epi($f$)不等式

$$F(\boldsymbol{x},\lambda)=(\boldsymbol{x}|\boldsymbol{u})+\alpha\lambda\geqslant\beta$$

恒成立. 因为λ可以充分大，所以$\alpha>0$.

因为 $(\boldsymbol{x},f(\boldsymbol{x}))\in$epi($f$) （设$\boldsymbol{x}\in$dom($f$)），所以

$$(\boldsymbol{x}|\boldsymbol{u})+\alpha f(\boldsymbol{x})\geqslant\beta. \tag{3.7-9}$$

因为点$(\boldsymbol{x}^{(0)},f(\boldsymbol{x}^{(0)}))\in H$，故

$$(\boldsymbol{x}^{(0)}|\boldsymbol{u})+\alpha f(\boldsymbol{x}^{(0)})=\beta. \tag{3.7-10}$$

由式(3.7-9)与(3.7-10)得

$$(\boldsymbol{x}|\boldsymbol{u})+\alpha f(\boldsymbol{x})\geqslant(\boldsymbol{x}^{(0)}|\boldsymbol{u})+\alpha f(\boldsymbol{x}^{(0)}).$$

由此可得

$$f(\boldsymbol{x})\geqslant f(\boldsymbol{x}^{(0)})+\left(\boldsymbol{x}-\boldsymbol{x}^{(0)}\middle|-\frac{1}{\alpha}\boldsymbol{u}\right)$$

当$\boldsymbol{x}\overline{\in}$dom($f$)时，上式显然也成立. 因而$f(\boldsymbol{x})$在$\boldsymbol{x}^{(0)}$是次可微

的，而且$-\frac{1}{\alpha}\boldsymbol{u}\in\partial f(\boldsymbol{x}^{(0)})$.

次梯度的几何解释. 由定理3.7.4的证明可知，$f(\boldsymbol{x})$在点$\boldsymbol{x}^{(0)}$的一个次梯度$\boldsymbol{u}$对应于上图epi(f)在点($\boldsymbol{x}^{(0)}$, $f(\boldsymbol{x}^{(0)})$)的一张非垂直的闭的支撑超平面H：

$$(\boldsymbol{x}|\boldsymbol{u})-\lambda=(\boldsymbol{x}^{(0)}|\boldsymbol{u})-f(\boldsymbol{x}^{(0)}). \qquad (3.7\text{-}11)$$

$(\boldsymbol{u}, -1)$是支撑超平面H的法矢. 单元可微函数$f(x)$在点$x^{(0)}$的次梯度就是切线H的斜率k. 若$f(\boldsymbol{x})$在点$\boldsymbol{x}^{(0)}$的次梯度不唯一，则上图epi(f)在点($\boldsymbol{x}^{(0)}$, $f(\boldsymbol{x}^{(0)})$)的支撑超平面也不唯一（参看图3-4）.

由定理3.7.3可以看出：次梯度是梯度概念的推广. 与此相关的是：支撑超平面可以看作多元函数微分学中所讨论过的空间曲面切平面的推广.

引理 设$f:E\to\mathscr{R}\cup\{+\infty\}$为正常凸函数，

$$\boldsymbol{x}^{(0)}\in\text{int}(\text{dom}(f)),$$

$$H:\ F(\boldsymbol{x}, \lambda)=(\boldsymbol{x}|\boldsymbol{u})+\alpha\lambda=\beta$$

为在点($\boldsymbol{x}^{(0)}$, $f(\boldsymbol{x}^{(0)})$)对epi(f)的一张闭的支撑超平面，则H是非垂直的，即$\alpha\neq 0$.

证 设$\alpha=0$. 因为点($\boldsymbol{x}^{(0)}$, $f(\boldsymbol{x}^{(0)})$)$\in H$，所以$(\boldsymbol{x}^{(0)}|\boldsymbol{u})=\beta$.

不妨设$F(\text{epi}(f))\geqslant\beta$. 当$\boldsymbol{x}\in\text{dom}(f)$时，点$(\boldsymbol{x}, f(\boldsymbol{x}))\in\text{epi}(f)$，因而得$(\boldsymbol{x}|\boldsymbol{u})\geqslant\beta$. 于是

$$(\boldsymbol{x}-\boldsymbol{x}^{(0)}|\boldsymbol{u})\geqslant 0, \quad \forall\boldsymbol{x}\in\text{dom}(f).$$

设$\boldsymbol{y}\in E$. 因为$\boldsymbol{x}^{(0)}\in\text{int}(\text{dom}(f))$，所以存在$\varepsilon>0$，使

$$\boldsymbol{x}^{(0)}+\varepsilon\boldsymbol{y}\in\text{dom}(f), \quad \boldsymbol{x}^{(0)}-\varepsilon\boldsymbol{y}\in\text{dom}(f).$$

于是得

$$\varepsilon(\boldsymbol{y}|\boldsymbol{u})\geqslant 0, \quad -\varepsilon(\boldsymbol{y}|\boldsymbol{u})\geqslant 0.$$

因而$(\boldsymbol{y}|\boldsymbol{u})=0$，$\forall\boldsymbol{y}\in E$. 表明$\boldsymbol{u}$为零值函数，与$H$为超平面矛

盾，所以$a\neq 0$.

定理 3.7.5 设f：$E\to\mathscr{R}\cup\{+\infty\}$为正常凸函数.

1° 若$f(\boldsymbol{x})$在某一点$\boldsymbol{x}^{(0)}\in\mathrm{dom}(f)$连续，则在有效域$\mathrm{dom}(f)$内$f(\boldsymbol{x})$是次可微的；

2° 若$E=\mathscr{R}^n$，则在有效域$\mathrm{dom}(f)$的相对内部$f(\boldsymbol{x})$是次可微的.

证 1° 因为$f(\boldsymbol{x})$在点$\boldsymbol{x}^{(0)}$连续，所以存在$\boldsymbol{x}^{(0)}$的一个邻域U与一个正数k，使$f(U)\leqslant k$. 于是当$\boldsymbol{x}\in U,\lambda\geqslant k$时，$(\boldsymbol{x},\lambda)\in\mathrm{epi}(f)$，因而$\mathrm{epi}(f)$的内部非空，$\mathrm{epi}(f)$是一个凸体. 另外还有$(\boldsymbol{x}^{(0)}, f(\boldsymbol{x}^{(0)}))\in\mathrm{bd}(\mathrm{epi}(f))$，所以在点$(\boldsymbol{x}^{(0)}, f(\boldsymbol{x}^{(0)}))$存在一张对$\mathrm{epi}(f)$的非平凡的闭的支撑超平面$H$（由定理2.2.3的推论）. 因为$\boldsymbol{x}^{(0)}\in\mathrm{int}(\mathrm{dom}(f))$，所以$H$是非垂直的（根据引理）. 根据定理3.7.4 $f(\boldsymbol{x})$在$\boldsymbol{x}^{(0)}$是次可微的.

根据定理3.4.6 $f(\boldsymbol{x})$在$\mathrm{dom}(f)$内每一点都是连续的. 按照上面的证明，$f(\boldsymbol{x})$在$\mathrm{dom}(f)$内每一点都是次可微的.

2° 设$\boldsymbol{x}^{(0)}\in\mathrm{ri}(\mathrm{dom}(f))$.显然$(\boldsymbol{x}^{(0)}, f(\boldsymbol{x}^{(0)}))\in\mathrm{rb}(\mathrm{dpi}(f))$. 于是在点$(\boldsymbol{x}^{(0)}, f(\boldsymbol{x}^{(0)}))$存在一张对$\mathrm{epi}(f)$的非平凡的闭的支撑超平面$H$（根据定理2.2.3的推论）. 用反证法可以证明超平面H是非垂直的. 推据定理3.7.4 $f(\boldsymbol{x})$在点$\boldsymbol{x}^{(0)}$是次可微的.由$\boldsymbol{x}^{(0)}$的任意性可得：在有效域$\mathrm{dom}(f)$的相对内部$f(\boldsymbol{x})$是次可微的.

推论 $\mathscr{R}^n$上的实凸函数$f(\boldsymbol{x})$在$\mathscr{R}^n$上是次可微的.

（因为$\mathrm{ri}(\mathrm{dom}(f))=\mathscr{R}^n$）

例2 设$C\subset E$为非空凸集，$\delta_C(\boldsymbol{x})$为C的指示函数.

1° 若$\boldsymbol{x}^{(0)}\in C$，则$\delta_C(\boldsymbol{x})$在点$\boldsymbol{x}^{(0)}$的次微分$\partial\delta_C(\boldsymbol{x}^{(0)})$是包含点$\boldsymbol{0}$的一个凸锥；

2° 若$\boldsymbol{x}^{(0)}\in\mathrm{int}(C)$，则$\partial\delta_C(\boldsymbol{x}^{(0)})=\{\boldsymbol{0}\}$.

证 1° 根据次微分的定义，$\boldsymbol{u}\in\partial\delta_C(\boldsymbol{x}^{(0)})$等价于

$$\delta_C(\boldsymbol{x})\geqslant\delta_C(\boldsymbol{x}^{(0)})+(\boldsymbol{x}-\boldsymbol{x}^{(0)}|\boldsymbol{u}),\quad\forall\,\boldsymbol{x}\in E.$$

若$x^{(0)}\in C$，则$\delta_C(x^{0})=0$，而且$\forall x\in C$均有$\delta_C(x)=0$，所以上式等价于

$$(x-x^{0}\,|\,u)\leqslant 0,\quad \forall x\in C.$$

因而

$$\partial\delta_C(x^{0})=\{u\,\big|\,(x-x^{0}\,|\,u)\leqslant 0,\ \forall x\in C\}.$$

容易验证上面的集合$\partial\delta_C(x^{0})$满足凸锥的充分必要条件（定理1.6.1），而且$\mathbf{0}\in\partial\delta_C(x^{(0)})$.

2° 若$x^{(0)}\in \mathrm{int}(C)$，则$\delta_C(x)$在$x^{0}$是可微的，而且$\delta_C(x)$在$x^{(0)}$的Fr′echet导数为$\mathbf{0}$. 根据定理3.7.3得

$$\partial\delta_C(x^{0})=\{\mathbf{0}\}.$$

*__例3__ 设$g:\mathscr{R}^n\to\mathscr{R}$为凸函数，而且满足Slater条件，即存在$x\in\mathscr{R}^n$使$g(x)<0$. 再设$x^{(0)}\in\mathscr{R}^n$使$g(x^{(0)})=0$，而

$$C=\{x\,|\,g(x)\leqslant 0\},$$
$$K=\{x\,|\,g'(x^{(0)};\ x)<0\},$$

则

$$\partial\delta_C(x^{(0)})=K^{0}. \tag{3.7-12}$$

证 因为$g(x)$是$\mathscr{R}^n$上的实凸函数，所以是连续的（根据定理3.4.8的推论1）而且是次可微的（根据定理3.7.5的推论）.

因为$g(x)$在$\mathscr{R}^n$上是连续的，所以C是闭的，即$C=\overline{C}$. 因为$g(x)$满足Slater条件，所以

$$\mathrm{int}(C)=\{x\,|\,g(x)<0\}\neq\varPhi.$$

根据定理1.4.5得

$$\overline{\mathrm{int}(C)}=\overline{C}=C. \tag{3.7-13}$$

$u\in\partial\delta_C(x^{0})$即

$$\delta_C(x)\geqslant\delta_C(x^{(0)})+(x-x^{(0)}\,|\,u),\quad \forall x\in\mathscr{R}^n.$$

因为$x^{(0)}\in C$，而且

$$\delta_C(x)=\begin{cases}0, & x\in C;\\ +\infty, & x\overline{\in}C.\end{cases}$$

所以上面的不等式等价于

$$(\boldsymbol{x}-\boldsymbol{x}^{0} \mid \boldsymbol{u})\leqslant 0\ ,\quad \forall \boldsymbol{x}\in C. \tag{3.7-14}$$

而根据$\boldsymbol{u}$的连续性与式(3.7-13)，式(3.7-14)等价于

$$(\boldsymbol{x}-\boldsymbol{x}^{(0)} \mid \boldsymbol{u})\leqslant 0\ ,\quad \forall \boldsymbol{x}\in \mathrm{int}(C). \tag{3.7-15}$$

因而$\boldsymbol{u}\in\partial\delta_C(\boldsymbol{x}^{(0)})$等价于不等式(3.7-15).

若$\boldsymbol{x}\in\mathrm{int}(C)$，$\lambda>0$，则

$$\begin{aligned} g'(\boldsymbol{x}^{0};\ \lambda(\boldsymbol{x}-\boldsymbol{x}^{0})) &= \lambda g'(\boldsymbol{x}^{0};\ \boldsymbol{x}-\boldsymbol{x}^{0}) \\ &= \lambda \lim_{t\to 0+}\frac{g(\boldsymbol{x}^{0}+t(\boldsymbol{x}-\boldsymbol{x}^{0}))-g(\boldsymbol{x}^{(0)})}{t} \\ &= \lambda \lim_{t\to 0+}\frac{g(t\boldsymbol{x}+(1-t)\boldsymbol{x}^{(0)})}{t} \\ &\leqslant \lambda \lim_{t\to 0+}\frac{tg(\boldsymbol{x})+(1-t)g(\boldsymbol{x}^{(0)})}{t} \\ &= \lambda g(\boldsymbol{x})<0\ , \end{aligned}$$

因而$\lambda(\boldsymbol{x}-\boldsymbol{x}^{0})\in K$. 于是得

$$\mathscr{R}_{+}(\mathrm{int}(C)-\boldsymbol{x}^{(0)})\subset K.$$

反之，若$\boldsymbol{y}\in K$，即$g'(\boldsymbol{x}^{0};\ \boldsymbol{y})<0$，则存在$t>0$使

$$\frac{g(\boldsymbol{x}^{0}+t\boldsymbol{y})-g(\boldsymbol{x}^{0})}{t}=\frac{g(\boldsymbol{x}^{0}+t\boldsymbol{y})}{t}<0\ ,$$

因而$g(\boldsymbol{x}^{0}+t\boldsymbol{y})<0$. 于是$\boldsymbol{x}^{(0)}+t\boldsymbol{y}\in\mathrm{int}(C)$. 由此得

$$\boldsymbol{y}\in\frac{1}{t}(\mathrm{int}(C)-\boldsymbol{x}^{(0)}),$$

因而

$$K\subset\mathscr{R}_{+}(\mathrm{int}(C)-\boldsymbol{x}^{(0)}),$$

所以

$$K=\mathscr{R}_{+}(\mathrm{int}(C)-\boldsymbol{x}^{(0)}).$$

设$\boldsymbol{y}\in K$，则存在$\lambda\in\mathscr{R}_{+}$，即$\lambda>0$，$\boldsymbol{x}\in\mathrm{int}(C)$，使

$$\boldsymbol{y}=\lambda(\boldsymbol{x}-\boldsymbol{x}^{0}),$$

因而，若式(3.7-15)成立，则

$$(\boldsymbol{y}|\boldsymbol{u})\leqslant 0,\quad \forall \boldsymbol{y}\in K. \tag{3.7-16}$$

若上式成立，也可推得不等式(3.7－15)．因而这两个不等式等价．于是得

$$\boldsymbol{u}\in\partial\delta_C(\boldsymbol{x}^0)\Longleftrightarrow\boldsymbol{u}\in K^0.$$

因而

$$\partial\delta_C(\boldsymbol{x}^{(0)})=K^0$$

定理 3.7.6 设$f:E\to\overline{\mathscr{R}}$为凸函数，$\boldsymbol{x}^{(0)}\in E$，$\partial f(\boldsymbol{x}^{(0)})\neq\Phi$，$\lambda>0$，则$\lambda f(\boldsymbol{x})$在$\boldsymbol{x}^0$的次微分不空，而且

$$\partial(\lambda f)(\boldsymbol{x}^{(0)})=\lambda\partial f(\boldsymbol{x}^{(0)}).$$

（由次梯度的定义直接可得）

定理 3.7.7 设f_1，$f_2:E\to\mathscr{R}\cup\{+\infty\}$为正常凸函数，则

$$\partial f_1(\boldsymbol{x})+\partial f_2(\boldsymbol{x})\subset\partial(f_1+f_2)(\boldsymbol{x}),\quad\forall\boldsymbol{x}\in E. \tag{3.7-17}$$

证 设$\boldsymbol{x}\in E$，$\boldsymbol{u}\in\partial f_1(x)$，$\boldsymbol{v}\in\partial f_2(\boldsymbol{x})$，即

$$f_1(\boldsymbol{y})\geqslant f_1(\boldsymbol{x})+(\boldsymbol{y}-\boldsymbol{x}|\boldsymbol{u}),\quad\forall\boldsymbol{y}\in E;$$

$$f_2(\boldsymbol{y})\geqslant f_2(\boldsymbol{x})+(\boldsymbol{y}-\boldsymbol{x}|\boldsymbol{v}),\quad\forall\boldsymbol{y}\in E.$$

将上面二式相加得

$$(f_1+f_2)(\boldsymbol{y})\geqslant(f_1+f_2)(\boldsymbol{x})+(\boldsymbol{y}-\boldsymbol{x}|\boldsymbol{u}+\boldsymbol{v}),\quad\forall\boldsymbol{y}\in E.$$

因此 $\boldsymbol{u}+\boldsymbol{v}\in\partial(f_1+f_2)(\boldsymbol{x})$，所以

$$\partial f_1(x)+\partial f_2(x)\subset\partial(f_1+f_2)(x),\quad\forall\boldsymbol{x}\in E.$$

注 当$\partial f_1(\boldsymbol{x})=\Phi$，$\partial f_2(\boldsymbol{x})=\Phi$时$\partial(f_1+f_2)(\boldsymbol{x})$可能不是空集（见例 4），因而式(3.7-17)的反包含关系不成立．

例 4 设有两个正常凸函数f_1，f_2如下：

$$f_1(x)=\begin{cases}+\infty, & x<0;\\ -\sqrt{x}, & x\geqslant 0.\end{cases}$$

$$f_2(x)=\begin{cases}-\sqrt{-x}, & x\leqslant 0;\\ +\infty, & x>0.\end{cases}$$

试确定次微分$\partial f_1(0)$，$\partial f_2(0)$与$\partial(f_1+f_2)(0)$．

解 因为对于任一实数a，不等式

$$f_1(x)\geqslant f_1(0)+ax,\quad \forall x\in\mathscr{R}$$

都不成立，所以次微分$\partial f_1(0)=\Phi$. 同样，次微分$\partial f_2(0)=\Phi$. 因为

$$(f_1+f_2)(x)=\begin{cases}0, & x=0;\\ +\infty, & x\neq 0.\end{cases}$$

对于任何实数a，不等式

$$(f_1+f_2)(x)\geqslant(f_1+f_2)(0)+ax,\quad \forall x\in\mathscr{R}$$

都成立，所以 $\partial(f_1+f_2)(0)=\mathscr{R}$.

例 4 表明下面的等式

$$\partial(f_1+f_2)(x)=\partial f_1(x)+\partial f_2(x)$$

在一般条件下不成立，仅在一定条件下成立.

定理 3.7.8 设f_1, f_2: $E\to\mathscr{R}\cup\{+\infty\}$为正常凸函数，$\boldsymbol{x}^{(1)}\in\mathrm{dom}(f_1)\cap\mathrm{dom}(f_2)$，$f_1(\boldsymbol{x})$在$\boldsymbol{x}^{(1)}$连续，则对于$\mathrm{dom}(f_1)\cap\mathrm{dom}(f_2)$的每一点$\boldsymbol{x}$都有

$$\partial(f_1+f_2)(\boldsymbol{x})=\partial f_1(\boldsymbol{x})+\partial f_2(\boldsymbol{x}). \tag{3.7-18}$$

***证** 设$\boldsymbol{x}^{(0)}\in\mathrm{dom}(f_1)\cap\mathrm{dom}(f_2)$. 先证明

$$\partial(f_1+f_2)(\boldsymbol{x}^{(0)})\subset\partial f_1(\boldsymbol{x}^{(0)})+\partial f_2(\boldsymbol{x}^{(0)}).$$

设 $\boldsymbol{u}\in\partial(f_1+f_2)(\boldsymbol{x}^{(0)})$. 令$g_1$, g_2: $E\to\mathscr{R}\cup\{+\infty\}$如下

$$g_1(\boldsymbol{x})=f_1(\boldsymbol{x}+\boldsymbol{x}^{(0)})-f_1(\boldsymbol{x}^{(0)})-(\boldsymbol{x}\mid\boldsymbol{u}),\quad \forall \boldsymbol{x}\in E;$$

$$g_2(\boldsymbol{x})=f_2(\boldsymbol{x}+\boldsymbol{x}^{(0)})-f_2(\boldsymbol{x}^{(0)}),\quad \forall \boldsymbol{x}\in E.$$

可以验证g_1, g_2在E上都是正常凸函数，而且具有以下性质：

$$\mathrm{dom}(g_1)=\mathrm{dom}(f_1)-\boldsymbol{x}^{(0)},$$

$$\mathrm{dom}(g_2)=\mathrm{dom}(f_2)-\boldsymbol{x}^{(0)},$$

$$g_1(\boldsymbol{x}(+g_2(\boldsymbol{x})\geqslant 0,\quad \forall \boldsymbol{x}\in E, \tag{3.7-19}$$

$$g_1(0)=g_2(0)=0.$$

因为$f_1(\boldsymbol{x})$在点$\boldsymbol{x}^{1}$是连续的，而$\boldsymbol{u}\in E'$是连续的，所以$g_1(\boldsymbol{x})$在点$\boldsymbol{x}^{(2)}=\boldsymbol{x}^{(1)}-\boldsymbol{x}^{(0)}$是连续的.

再令

$$A=\{(\boldsymbol{x},\ \lambda)\mid g_1(\boldsymbol{x})\leqslant\lambda\},$$

$$B=\{(\boldsymbol{x},\ \lambda\mid\lambda\leqslant-g_2(\boldsymbol{x})\}.$$

$A=\mathrm{epi}(g_1)$是凸集，B是凹函数$-g_2(\boldsymbol{x})$的下图，因而也是凸集。显然，A，B均非空($(0,\ 0)\in A\cap B$. 因为$g_1(\boldsymbol{x})$在$\boldsymbol{x}^{(2)}$是连续的，所以

$$(\boldsymbol{x}^{(2)},\ g_1(\boldsymbol{x}^{(2)})+1)\in\mathrm{int}(A). \qquad (3.7\text{-}20)$$

因而$\mathrm{int}(A)\neq\Phi$. 用反证法可以证明

$$\mathrm{int}(A)\cap B=\Phi.$$

倘若不然，存在$(\boldsymbol{y},\ \alpha)\in\mathrm{int}(A)\cap B$，则存在$\varepsilon>0$，使$(\boldsymbol{y},\alpha-\varepsilon)\in A$，因而

$$g_1(\boldsymbol{y})\leqslant\alpha-\varepsilon<\alpha.$$

由$(\boldsymbol{y},\ \alpha)\in B$得$\alpha\leqslant-g_2(\boldsymbol{y})$. 于是得不等式

$$g_1(\boldsymbol{y})+g_2(\boldsymbol{y})<0.$$

上式与不等式(3.7-19)矛盾，因而有

$$\mathrm{int}(A)\cap B=\Phi.$$

根据定理2.2.2 $E\times\mathscr{R}$中存在一张分离A与B的闭超平面H.不妨将超平面H写作

$$(\boldsymbol{x}\mid\boldsymbol{v})+\alpha\lambda=\beta,$$

其中$(\boldsymbol{v},\ \alpha)\in(E\times\mathscr{R})'$，而且$(\boldsymbol{v},\ \alpha)\neq(0,\ 0)$.

不妨设

$$(\boldsymbol{x}\mid\boldsymbol{v})+\alpha\lambda\geqslant\beta,\ \forall(\boldsymbol{x},\ \lambda)\in A;$$

$$(\boldsymbol{x}\mid\boldsymbol{v})+\alpha\lambda\leqslant\beta,\ \forall(\boldsymbol{x},\ \lambda)\in B.$$

因为$(0,\ 0)\in A\cap B$，所以$\beta=0$. 因为$\forall\lambda>0$，$(0,\ \lambda)\in A$,所以$\alpha\geqslant0$. 可以证明$\alpha\neq0$. 倘若$\alpha=0$，则

$$(\boldsymbol{x}\mid\boldsymbol{v})\geqslant0,\ \forall\boldsymbol{x}\in\mathrm{dom}(g_1);$$

$$(\boldsymbol{x}\mid\boldsymbol{v})\geqslant0,\ \forall\boldsymbol{x}\in\mathrm{dom}(g_2).$$

因为$\boldsymbol{x}^{(2)}=\boldsymbol{x}^{(1)}-\boldsymbol{x}^{(0)}$属于$\mathrm{dom}(g_1)\cap\mathrm{dom}(g_2)$，所以$(\boldsymbol{x}^{(2)}\mid\boldsymbol{v})$

$=0$. 因而$(x^{(2)}, g(x^{(2)})+1)\in H$. 然而$(x^{(2)}, g(x^{(2)})+1)$是$A$的内点（由式(3.7-20)知）. 这与$H$分离$A$与$B$矛盾，所以$\alpha\neq 0$. 于是$\alpha>0$.

设 $w=\frac{1}{\alpha}v$, 则 $(x|w)=\frac{1}{\alpha}(x|v)$.

当$x\in \mathrm{dom}(g_1)$时，$(x, g_1(x))\in A$, 因而

$$(x|v)+\alpha g_1(x)\geqslant 0.$$

由此可得

$$g_1(x)\geqslant -(x|w).$$

当$x\overline{\in}\mathrm{dom}(g_1)$时，上式显然也成立. 再根据$g_1(x)$的定义得

$$f_1(x+x^{(0)})\geqslant f_1(x^{(0)})+(x|u-w),\quad \forall x\in D.$$

同理可得

$$f_2(x+x^{(0)})\geqslant f_2(x^{(0)})+(x|w),\quad \forall x\in E.$$

因而$u-w\in\partial f_1(x^{(0)})$, $w\in\partial f_2(x^{(0)})$. 于是得

$$u=(u-w)+w\in\partial f_1(x^{(0)})+\partial f_2(x^{(0)}),$$

所以

$$\partial(f_1+f_2)(x^{(0)})\subset\partial f_1(x^{(0)})+\partial f_2(x^{(0)}).$$

根据定理3.7.7上式的反包含关系也成立，于是

$$\partial(f_1+f_2)(x^{(0)})=\partial f_1(x^{(0)})+\partial f_2(x^{(0)})$$

对于有限维情形，定理3.7.8中的条件可以放宽.

定理 3.7.9 设$f_1, f_2: \mathscr{R}^n\to\mathscr{R}\cup\{+\infty\}$为正常凸函数，而且

$$\mathrm{ri}(\mathrm{dom}(f_1))\cap \mathrm{ri}(\mathrm{dom}(f_2))\neq\Phi,$$

则

$$\partial(f_1+f_2)(x)=\partial f_1(x)+\partial f_2(x),\quad \forall x\in\mathscr{R}^n.$$

证 （仿定理3.7.8的证明进行，略）

虽然定理3.7.9的条件比定理3.7.8的条件弱，可是它仍然只是充分的，不是必要的. 请见下例.

例 5 设f_1, f_2如下：

$$f_1(x)=\begin{cases}+\infty, & x<0;\\ e^x, & x\geqslant 0.\end{cases}$$

$$f_2(x)=\begin{cases}e^{-x}, & x\leqslant 0;\\ +\infty, & x<0.\end{cases}$$

试确定次微分$\partial f_1(x)$，$\partial f_2(x)$，$\partial(f_1+f_2)(x)$.

解 1° 当$x>0$时，$f_1(x)$可微，$f_1'(x)=e^x$，所以

$$\partial f_1(x)=\{e^x\}.$$

当$x<0$时，$f_1(x)=+\infty$，而$f_1(0)=1$，所以任何实数a都不是$f_1(x)$的次梯度.

在$x=0$处，对于任何实数$a\in(-\infty,\ 1]$，不等式

$$f_1(y)\geqslant f_1(0)+(y|a) \qquad \forall y\in\mathscr{R}$$

都成立，而当$a>1$时，上面的不等式不成立（见图3-5）. 因而$\partial f_1(0)=(-\infty,\ 1]$. 综上所述得

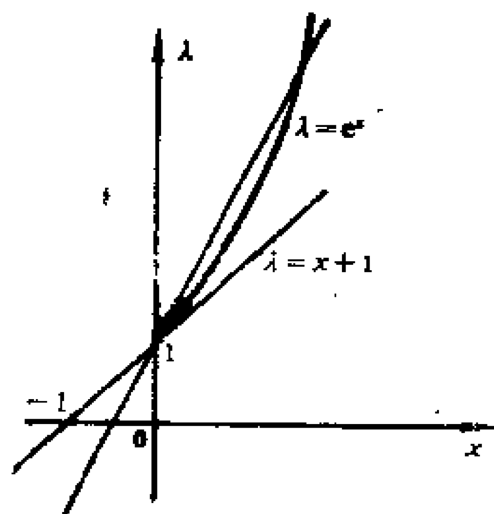


图 3-5

$$\partial f_1(x)=\begin{cases}\Phi, & x<0;\\ (-\infty,\ 1], & x=0;\\ \{e^x\}, & x>0.\end{cases}$$

2° 仿1°可得

$$\partial f_2(x)=\begin{cases}\{-e^{-x}\}, & x<0;\\ [-1,\ +\infty), & x=0;\\ \Phi, & x>0.\end{cases}$$

3° $f_1(x)$与$f_2(x)$的和如下：

$$(f_1+f_2)(x)=\begin{cases}2, & x=0;\\ +\infty, & x\neq 0.\end{cases}$$

$f_1(x)+f_2(x)$的次微分如下：

$$\partial(f_1+f_2)(x)=\begin{cases}\mathscr{R}, & x=0;\\ \Phi, & x\neq 0.\end{cases}$$

注 对于此例

$$\mathrm{ri}(\mathrm{dom}(f_1))\cap \mathrm{ri}(\mathrm{dom}(f_2))=\Phi.$$

定理3.7.9的条件没有满足，然而以下等式

$$\partial(f_1+f_2)(x)=\partial f_1(x)+\partial f_2(x),\quad \forall x\in\mathscr{R}$$

成立.

§3.8 Jensen不等式与Hadamard不等式

与凸函数有关的不等式很多，本节将研究两个重要的不等式，并介绍其应用.

3.8.1 Jensen不等式

在§3.3中所提到的不等式(3.3-4)称为Jensen不等式．Jensen不等式有多种形式．式(3.3-4)是一种特殊形式．下面研究积分形式的Jensen不等式.

定理 3.8.1（Jensen不等式） 设$f(x)$为区间I上的实值连续凸函数，g:〔α, β〕→ I 是逐段连续函数，至多具有有限多个第一类间断点，则以下不等式

$$f\left(\frac{1}{\beta-\alpha}\int_\alpha^\beta g(t)\mathrm{d}t\right)\leqslant\frac{1}{\beta-\alpha}\int_\alpha^\beta f(g(t))\mathrm{d}t \quad (3.8\text{-}1)$$

成立.

证 若$g(t)$在〔α, β〕上为常数，则式(3.8-1)显然成立．若$g(t)$不是常数，令

$$p=\frac{1}{\beta-\alpha}\int_\alpha^\beta g(t)\mathrm{d}t, \quad (3.8\text{-}2)$$

则$p\in\mathrm{int}(I)$．由$f(x)$的凸性可得：$f(x)$在点p的右导数$f'_+(p)$存在，而且以下不等式成立

$$f(x)\geqslant f(p)+f'_+(p)(x-p),\quad \forall x\in I.$$

在上式中令$x=g(t)$得

$$f(g(t))\geqslant f(p)+f'_+(p)(g(t)-p),\quad \forall t\in〔\alpha,\ \beta〕.$$

将上式两端在区间$[\alpha, \beta]$上积分 得

$$\int_\alpha^\beta f(g(t))\mathrm{d}t \geqslant f(p)(\beta-\alpha)+f'_+(p)\left[\int_\alpha^\beta g(t)\mathrm{d}t-p(\beta-\alpha)\right]$$

$$=f(p)(\beta-\alpha).$$

将上式两端除以$\beta-\alpha$即得不等式(3.8-1).

推论 若$f(x)$为区间I上的实值凸函数，$x_i\in I$，$i=1, \cdots, m$，$\lambda_i>0$，$\lambda_1+\cdots+\lambda_m=1$，则

$$f\left(\sum_{i=1}^m \lambda_i x_i\right)\leqslant \sum_{i=1}^m \lambda_i f(x_i). \qquad (3.8\text{-}3)$$

此不等式即式(3.3-4)，在§3.3中已用其它方法证得．若令$\alpha=0$，$\beta=1$，$t_0=\alpha$，$t_i=t_{i-1}+\lambda_i$，$g(t)=x_i$，$t\in(t_{i-1}, t_i]$，$i=1, \cdots, m$，则式(3.8-1)化为式(3.8-3).

注 若$f: I\to\mathscr{R}$为严格凸函数，$x_i\in I$，$i=1, \cdots, m$，x_i不完全相等，$\lambda_i>0$，并且$\lambda_1+\cdots+\lambda_m=1$，则以下严格不等式

$$\left(\sum_{i=1}^m \lambda_i x_i\right)<\sum_{i=1}^m \lambda_i f(x_i) \qquad (3.8\text{-}4)$$

成立.

（用数学归纳法可以证明）

3.8.2 Hadamard不等式

定理 3.8.2（Hadamard不等式） 设$f(x)$为区间$[a, b]$上的实值连续凸函数，则以下不等式

$$f\left(\frac{a+b}{2}\right)\leqslant\frac{1}{b-a}\int_a^b f(x)\mathrm{d}x\leqslant\frac{f(a)+f(b)}{2} \qquad (3.8\text{-}5)$$

成立.

证 1° 因为$f(x)$为凸函数，所以（根据定理3.6.1）有不等式

$$f(x)\leqslant\frac{b-x}{b-a}f(a)+\frac{x-a}{b-a}f(b), \quad \forall x\in[a, b].$$

将上式两端积分得

$$\int_a^b f(x)\mathrm{d}x \leqslant \frac{f(a)}{b-a}\int_a^b (b-x)\mathrm{d}x + \frac{f(b)}{b-a}\int_a^b (x-a)\mathrm{d}x$$

$$=(f(a)+f(b))\cdot\frac{b-a}{2},$$

两端除以$b-a$，得

$$\frac{1}{b-a}\int_a^b f(x)\mathrm{d}x \leqslant \frac{f(a)+f(b)}{2}. \qquad (3.8\text{-}6)$$

2° 在Jensen不等式(3.8-1)中，取$\alpha=a$，$\beta=b$，$g(t)=t$，于是得

$$f\left(\frac{1}{b-a}\int_a^b t\ \mathrm{d}t\right) \leqslant \frac{1}{b-a}\int_a^b f(t)\mathrm{d}t,$$

此即

$$f\left(\frac{a+b}{2}\right) \leqslant \frac{1}{b-a}\int_a^b f(x)\mathrm{d}x \qquad (3.8\text{-}7)$$

综合式(3.8-6)与(3.8-7)即得Hadamard不等式(3.8-5).

3.8.3 应用

算术平均-几何平均不等式是几何规划的基础. 作为Jensen不等式的应用，首先研究其基本形式.

定理 3.8.3（算术平均-几何平均不等式） 设a_1，…，$a_n>0$，$\lambda_i>0$，$\lambda_1+\cdots+\lambda_n=1$，则以下不等式

$$\prod_{i=1}^{n} a_i^{\lambda_i} \leqslant \sum_{i=1}^{n} \lambda_i a_i \qquad (3.8\text{-}8)$$

成立. 当且仅当所有的a_i全等时，不等式(3.8-8)化为等式.

证 1° 因为指数函数$\exp(x)=\mathrm{e}^x$是凸函数，根据Jensen不等式(3.8-3)得

$$\exp\left(\sum_{i=1}^{n}\lambda_i\ln a_i\right)\leqslant\sum_{i=1}^{n}\lambda_i\exp(\ln a_i),\qquad(3.8\text{-}9)$$

化简后即得不等式(3.8-8).

2° 若$a_1=a_2=\cdots=\cdots a_n$，则式(3.8-8)两端都等于$a_1$，所以式(3.8-8)化为等式.

因为指数函数exp(x)是严格凸函数，所以若a_i不完全相等，则不等式(3.8-3)应为严格不等式(3.8-4)，因而不等式(3.8-9)也是严格不等式. 于是式(3.8-8)也是严格不等式. 由此可知：若式(3.8-8)是等式，必定是所有a_i都相等.

推论 若$a_i>0$，$p_i>0$，$i=1$，…，n，则

$$a_1^{p_1}a_2^{p_2}\cdots a_n^{p_n}\leqslant\left(\frac{p_1a_1+p_2a_2+\cdots+p_na_n}{p}\right)^p,\qquad(3.8\text{-}10)$$

其中$p=p_1+p_2+\cdots+p_n$. 当且仅当所有的a_i都相等时上式化为等式.

证 令$\lambda_i=p_i/p$，$i=1$，…，n，则$\lambda_i>0$，并且$\lambda_1+\cdots+\lambda_n=1$. 于是不等式(3.8-8)成立. 用$p_i/p$代其中的$\lambda_i$，得不等式

$$\left(\prod_{i=1}^{n}a_i^{p_i}\right)^{\frac{1}{p}}\leqslant\sum_{i=1}^{n}\frac{p_i}{p}a_i,$$

将上式两端各取对数得不等式

$$\frac{1}{p}\ln\left(\prod_{i=1}^{n}a_i^{p_i}\right)\leqslant\ln\left(\sum_{i=1}^{n}\frac{p_i}{p}a_i\right).$$

由此可得

$$\ln\left(\prod_{i=1}^{n}a_i^{p_i}\right)\leqslant\ln\left(\sum_{i=1}^{n}\frac{p_i}{p}a_i\right)^p.$$

将上式两端去对数即得不等式(3.8-10).

作为Hadamard不等式的应用，证明下面的定理.

定理 3.8.4 设$f(x)$为闭区间$[a, b]$上的连续凸函数，则函数

$$F(x)=\int_a^x f(t)\mathrm{d}t-(x-a)f\left(\frac{x+a}{2}\right) \qquad (3.8\text{-}11)$$

在$[a, b]$上非减.

证 设$a\leqslant x_1<x_2\leqslant b$，则

$$F(x_2)=\int_a^{x_2} f(t)\mathrm{d}t-(x_2-a)f\left(\frac{x_2+a}{2}\right),$$

$$F(x_1)=\int_a^{x_1} f(t)\mathrm{d}t-(x_1-a)f\left(\frac{x_1+a}{2}\right).$$

将以上二式相减得

$$F(x_2)-F(x_1)=\int_{x_1}^{x_2} f(t)\mathrm{d}t-(x_2-a)f\left(\frac{x_2+a}{2}\right)$$

$$+(x_1-a)f\left(\frac{x_1+a}{2}\right).$$

根据Hadamard不等式(3.8-5)得

$$\int_{x_1}^{x_2} f(t)\mathrm{d}t\geqslant(x_2-x_1)f\left(\frac{x_1+x_2}{2}\right).$$

根据$f(x)$的凸性得

$$f\left(\frac{x_2+a}{2}\right)=f\left(\frac{x_2-x_1}{x_2-a}\cdot\frac{x_2+x_1}{2}+\frac{x_1-a}{x_2-a}\cdot\frac{x_1+a}{2}\right)$$

$$\leqslant\frac{x_2-x_1}{x_2-a}f\left(\frac{x_2+x_1}{2}\right)+\frac{x_1-a}{x_2-a}f\left(\frac{x_1+a}{2}\right).$$

综合以上三式得

$$F(x_2)-F(x_1)\geqslant(x_2-x_1)f\left(\frac{x_1+x_2}{2}\right)-(x_2-x_1)f\left(\frac{x_2+x_1}{2}\right)$$

$$=0,$$

所以$F(x)$在$[a, b]$上非减.

在数学分析中有一个定积分性质（估值不等式）如下：

若$f(x)$在闭区间$[a, b]$上连续，M，m分别为$f(x)$在$[a, b]$上的最大值与最小值，则

$$m(b-a)\leqslant\int_a^b f(x)\mathrm{d}x\leqslant M(b-a),$$

亦即

$$m\leqslant\frac{1}{b-a}\int_a^b f(x)\mathrm{d}x\leqslant M. \qquad (3.8\text{-}12)$$

若$f(x)$为凸函数，则可用式(3.8-5)代替上式.

例 设Q由下式给出，试估计它的值.

$$Q=\frac{1}{4}\int_1^5\sqrt{x^4+1}\,\mathrm{d}x.$$

解法一 因为函数$f(x)=\sqrt{x^4+1}$在闭区间$[1, 5]$上的最小值与最大值（取四位小数）分别为

$$m=\sqrt{2}\approx 1.4142,$$

$$M=\sqrt{626}\approx 25.0200.$$

所以按照公式(3.8-12)得

$$1.4142\leqslant Q\leqslant 25.0200,$$

即

$$Q\in[1.4142,\ 25.0200].$$

解法二 因为

$$f'=\frac{2x^3}{\sqrt{x^4+1}},\qquad f''=\frac{2x^6+6x^2}{(x^4+1)^{3/2}}>0,$$

所以$f(x)=\sqrt{x^4+1}$在$[1, 5]$上是连续的凸函数.

$$f\left(\frac{1+5}{2}\right)=\sqrt{3^4+1}=\sqrt{82}\approx 9.0553,$$

$$\frac{1}{2}(f(1)+f(5))=\frac{1}{2}(\sqrt{2}+\sqrt{626})\approx 13.2172,$$

按Hadamard不等式(3.8-5)得

$$9.0553\leqslant Q\leqslant 13.2172,\quad 即\quad Q\in[9.0553,\ 13.2172].$$

比较 按估值不等式(3.8-12)所得区间的长度为

$$d_1=25.0200-1.4142=23.6058.$$

按Hadamard不等式(3.8-5)所得区间的长度为

$$d_2=13.2172-9.0553=4.1619.$$

$$d_2:\ d_1<1/5.$$

由此可见：用Hadamard不等式比用估值不等式要精确。

习 题

1. 设$C\subset V$，f：$C\to\overline{\mathscr{R}}$为凸函数，试证明有效域$\mathrm{dom}(f)$为凸集.

2. 设$f_1=x^2+1$，$f_2=e^{2x}$，$f(x)=\sup\limits_{i=1,2}\{f_i(x)\}$，$x\in\mathscr{R}$.

 1° 试描绘$f(x)$的图形；

 2° 试证明$\mathrm{epi}(f)$为凸集.

3. 设$A\subset V\times\mathscr{R}$为凸集. 试证明

$$f(\boldsymbol{x})=\sup\{\lambda\mid(\boldsymbol{x},\ \lambda)\in A\}$$

为凹函数，且$f(\boldsymbol{x})$的下图包含A.

*4. 设$f(\boldsymbol{x})$为线性拓扑空间E上的连续正常凸函数，$\lambda\in\mathscr{R}$，

$$A=\{\boldsymbol{x}\mid f(\boldsymbol{x})\leqslant\lambda\},\quad B=\{\boldsymbol{x}\mid f(\boldsymbol{x})<\lambda\}.$$

试证明：若$B\neq\Phi$，则$\mathrm{int}(A)=B$.

5. 设$f(\boldsymbol{x})$定义在凸集$D\subset\mathscr{R}^n$上，$\boldsymbol{x}'$，$\boldsymbol{x}''\in D$.

$$\phi(t)=f(t\boldsymbol{x}'+(1-t)\boldsymbol{x}''),\quad t\in[0,\ 1]$$

则

1° $f(\boldsymbol{x})$在凸集D上为凸函数的充要条件是：$\forall\,\boldsymbol{x}',\boldsymbol{x}''\in D$，单元函数$\phi(t)$在区间$[0,\ 1]$上均为凸函数；

2° 若$\forall x' \neq x''$，$\phi(t)$在$[0, 1]$上为严格凸函数，则$f(x)$在D上为严格凸函数.

6. 设函数φ: $\mathscr{R}^n \times \mathscr{R} \to \mathscr{R}$为连续函数，且$\forall t \in [a, b]$，函数$\varphi(x, t)$关于$x$是凸函数. 试证明函数

$$f(x) = \int_a^b \varphi(x, t)\mathrm{d}t, \quad x \in \mathscr{R}^n$$

为凸函数.

*7. 设$I \subset \mathscr{R}$为区间. 若$\forall x, x' \in I$均有

$$f\left(\frac{x+x'}{2}\right) \leqslant \frac{1}{2}[f(x)+f(x')]$$

则称$f(x)$为次凸函数（或中点凸函数）. 试证明：若$f(x)$为次凸函数而且连续，则$f(x)$为凸函数.

8. 设函数

$$f(x) = \begin{cases} x^2, & x < 0; \\ 2x+1, & x \geqslant 0. \end{cases}$$

1° 描绘$f(x)$的上图$\mathrm{epi}(f)$;

2° 求$f(x)$的下半连续包$\overline{f}(x)$;

3° 求$f(x)$的凸包$\mathrm{co}(f)$.

9. 设$I \subset \mathscr{R}$为一个区间，f: $I \to \mathscr{R}$为凸函数，$c \in \mathrm{int}(I)$. 试证明$f(x)$在点c的右导数$f'_+(c)$存在，而且以下不等式成立

$$f(x) \geqslant f(c) + f'_+(c)(x-c), \quad \forall x \in I.$$

10. 设

$$f(x) = \begin{cases} x^2-2x+1, & x \leqslant 0; \\ e^x, & x > 0. \end{cases}$$

试求$f'_-(x)$，$f'_+(x)$，$f'_-(0)$，$f'_+(0)$，$f'_-(1)$，$f'_+(1)$.

11. 设$f(x)$同上题. 试求

1° $f(x)$在点$x=0$的次微分$\partial f(0)$;

2° $f(x)$在任一点x的次微分.

12. 设$f(x)$同习题10，试写出$f(x)$在点$c=(0, 1)^T$对上图$\text{epi}(f)$的一张闭的支撑超平面. 在该点支撑超平面是否唯一?

*13. 设$f_1, f_2: \mathscr{R}^n \to \mathscr{R} \cup \{+\infty\}$均为正常凸函数，而且

$$\text{ri}(\text{dom}(f_1)) \cap \text{ri}(\text{dom}(f_2)) \neq \Phi.$$

试证明

$$\partial(f_1+f_2)(\boldsymbol{x}) = \partial f_1(\boldsymbol{x}) + \partial f_2(\boldsymbol{x}), \quad \forall \boldsymbol{x} \in \mathscr{R}^n.$$

14. 设$f: I \to \mathscr{R}$为严格凸函数，$x_i \in I$, $i=1, \cdots, n$, x_i不完全相等，$\lambda_i > 0$，并且$\lambda_1 + \cdots + \lambda_n = 1$，试证明以下严格不等式成立

$$f\left(\sum_{i=1}^{n} \lambda_i x_i\right) < \sum_{i=1}^{n} \lambda_i f(x_i).$$

15. 设$f(x) = (x^2+1)^{3/2} e^{2x}$，试用Hadamard不等式估计定积分$\int_0^1 f(x) dx$的值.

第四章 凸规划

本章首先介绍有关凸规划的一些基本概念，然后讨论凸函数的极值与凸规划最优解的充分必要条件.

§4.1 基本概念

定义 4.1.1 若$S\subset\mathscr{R}^n$为凸集，$f(\boldsymbol{x})$为S上的凸函数，则下述问题

$$(\mathrm{CP}_1)\qquad \min_{\boldsymbol{x}\in S}\ z=f(\boldsymbol{x}) \qquad (4.1\text{-}1)$$

称为**凸规划**（convex programming）. 若$f(\boldsymbol{x})$为严格凸函数，则(CP_1)称为严格凸规划.

令
$$F(\boldsymbol{x})=\begin{cases} f(\boldsymbol{x}), & \boldsymbol{x}\in S;\\ +\infty, & \boldsymbol{x}\overline{\in} S.\end{cases}$$

则S上的凸函数$f(\boldsymbol{x})$就开拓为$\mathscr{R}^n$上的正常凸函数$F(\boldsymbol{x})$. 因而凸规划（CP_1）等价于无约束问题

$$\min F(\boldsymbol{x}). \qquad (4.1\text{-}2)$$

若在$\mathscr{R}^n$上$f(\boldsymbol{x})$为凸函数，$g_j(\boldsymbol{x})$为凹函数，$h_i(\boldsymbol{x})$为仿射函数，$j=1,\cdots,p$，$i=1,\cdots,m$，则数学规划

$$\min\ z=f(\boldsymbol{x}), \qquad (4.1\text{-}3)$$

$$(\mathrm{MP})\qquad \text{s.t.}\quad h_i(\boldsymbol{x})=0,\ i=1,\cdots,m \qquad (4.1\text{-}4)$$

$$g_j(\boldsymbol{x})\geqslant 0,\ j=1,\cdots,p \qquad (4.1\text{-}5)$$

就是凸规划（容易验证(MP)的可行域是凸集）.

线性规划是凸规划的一个特例.

定义 4.1.2 设$S\subset\mathscr{R}^n$为(MP)的可行域，$x^*\in S$. 若

$$f(x)\geqslant f(x^*),\quad \forall x\in S, \tag{4.1-6}$$

则称x^*为(MP)的**总体最优解**(或**总体极小点**).

定义 4.1.3 设$S\subset\mathscr{R}$为(MP)的可行域，$x^0\in S$. 若存在x^0的一个邻域U，使

$$f(x)\geqslant f(x^0)\qquad \forall x\in S\cap U, \tag{4.1-7}$$

则称x^0为(MP)的**局部最优解**(或**局部极小点**).

若式(4.1-6)、(4.1-7)中的不等式为严格不等式，则x^*、$x^{(0)}$分别称为$f(x)$的**严格总体极小点与严格局部极小点**.

仿此可以给出$f(x)$在S上的总体极大点、局部极大点的定义.

显然，$f(x)$的总体极小点也是$f(x)$的局部极小点. $f(x)$的总体极大点也是$f(x)$的局部极大点.

定义 4.1.4 设$x^{(1)}$是(MP)的可行点. 若$g_j(x^{(1)})=0$，则称约束$g_j(x)\geqslant 0$在点$x^{(1)}$是**起作用约束或有效约束**(active constraint).

在可行点$x^{(1)}$，等式约束$h_i(x)=0$都是起作用约束.

定义 4.1.5 设$h_i(x)$，$g_j(x)\in C^1$，$i=1,\cdots,m$，$j=1,\cdots,p$，$x^{(1)}$是(MP)的可行点. 若在$x^{(1)}$起作用约束的梯度

$$\nabla h_i(x^{(1)}),\quad i=1,\cdots,m,$$

$$\nabla g_j(x)^{(1)}),\quad j\in I=\{j\mid(g_jx^{(1)})=0\},$$

线性无关，则称$x^{(1)}$为约束(4.1-4)与(4.1-5)的**正则点**(regular point).

§4.2 凸函数的极小点

定理 4.2.1 若$S\subset\mathscr{R}^n$为凸集，$f(x)$为定义在S上的实值凸函数，则$f(x)$的任一局部极小点$x^{(0)}$必为$f(x)$的总体极小点.

证 设$x^{(0)}$为$f(x)$的局部极小点，x为S中任一点. 必存在

$x^{(0)}$的邻域U，对于$S\cap U$中任一点y都有$f(y)\geqslant f(x^{(0)})$.

当$\lambda\in(0,\ 1)$充分接近于1时，$\lambda x^{(0)}+(1-\lambda)x\in S\cap U$因而

$$f(\lambda x^{(0)}+(1-\lambda)x)\geqslant f(x^{(0)}). \qquad (4.2\text{-}1)$$

再由$f(x)$的凸性得

$$f(\lambda x^{(0)}+(1-\lambda)x)\leqslant\lambda f(x^{(0)})+(1-\lambda)f(x). \qquad (4.2\text{-}2)$$

由式(4.2-1)、式(4.2-2)得

$$f(x^{(0)})\leqslant\lambda f(x^{(0)})+(1-\lambda)f(x).$$

由此可得$f(x)\geqslant f(x^{(0)})$. 因为x是任意的，所以$x^{(0)}$为$f(x)$的总体极小点.

定理 4.2.2 若$S\subset\mathscr{R}$为凸集，$f(x)$为定义在S上的实值凸函数，则$f(x)$的全部极小点组成凸集.

证 设x^*为$f(x)$的任一极小点，极小值为$f^*=f(x^*)$. $f(x)$的全部极小点组成水平集

$$L(f,\ f^*)=\{x\mid x\in S,\ f(x)\leqslant f^*\} \qquad (4.2\text{-}3)$$

而水平集$L(f,\ f^*)$是凸集，所以$f(x)$的全部极小点组成凸集.

定理 4.2.3 设$f(x)$为凸集$S\subset\mathscr{R}^n$上的实值严格凸函数.若极小点存在，则必唯一.

证 倘若不然，设$x^{(1)}$，$x^{(2)}$都是$f(x)$的极小点，且$x^{(1)}\neq x^{(2)}$. 取$\alpha\in(0,\ 1)$，则$\alpha x^{(1)}+(1-\alpha)x^{(2)}\in S$. 由$f(x)$的严格凸性得

$$\begin{aligned}f(\alpha x^{(1)}+(1-\alpha)x^{(2)})&<\alpha f(x^{(1)})+(1-\alpha)f(x^{(2)})\\&=f(x^{(1)}),\end{aligned}$$

这与$x^{(1)}$为$f(x)$的极小点（即总体极小点）矛盾. 所以$x^{(1)}=x^{(2)}$，即极小点唯一.

根据凸规划的定义，也可以将以上三个定理概括如下：**凸规划的局部最优解就是总体最优解；凸规划的最优解组成凸集；严格凸规划的最优解若存在则必唯一.**

现在研究凸规划在几何规划中的应用。下面的数学规划称为**正项几何规划**（posynomial geometric programming）。

$$\min\ y_0(x)=\sum_{j=1}^{T_0} c_{0j}f_{0j}(x)$$

$$\text{(GP)}\quad \text{s.t.}\quad y_m(x)=\sum_{j=1}^{T_m} c_{mj}f_{mj}(x)\leqslant 1,\ m=1,\ \cdots,\ M$$

$$x=(x_1,\ x_2,\ \cdots,\ x_n)^T>0,$$

其中

$$f_{mj}(x)=\prod_{i=1}^{n} x_i^{a_{mij}},\ m=0,\ 1,\ \cdots,\ M;\ j=1,\ \cdots,\ T_m,$$

所有系数c_{mj}均为正实数；而指数a_{mij}为实数。

正向几何规划不一定是凸规划，然而下述命题是成立的。

定理 4.2.4 正项几何规划(GP)的任一局部最优解都是它的总体最优解。

证 令

$$x_i=\exp(u_i),\ i=1,\ \cdots,\ n,\ u_i\in\mathscr{R},$$

$$Y_m(u)=y_m(x)$$

$$=\sum_{j=1}^{T_m} c_{mj}\exp\left(\sum_{i=1}^{n} a_{mij}u_i\right),\ m=0,\ 1,\ \cdots,\ M,$$

则正项几何规划(GP)化为

$$\text{(CP}_2\text{)}\quad \min\ Y_0(u),$$

$$\text{s.t.}\quad Y_m(u)\leqslant 1,\ m=1,\ \cdots,\ M,$$

其中$u=(u_1,\ \cdots,\ u_n)^T$。

因为指数函数$e^t=\exp(t)$是严格凸函数，所以对于任一$\lambda\in(0,$

1）均有

$$
\begin{aligned}
&\exp\left(\sum_{j=1}^{n} a_{mij}(\lambda u_i+(1-\lambda)v_i)\right)\\
&=\exp\left(\lambda\sum_{i=1}^{n} a_{mij}u_i+(1-\lambda)\sum_{i=1}^{n} a_{mij}v_i\right)\\
&<\lambda\exp\left(\sum_{i=1}^{n} a_{mij}u_i\right)\\
&\quad+(1-\lambda)\exp\left(\sum_{i=1}^{n} a_{mij}v_i\right).
\end{aligned}
$$

因而$\exp\left(\sum_{i=1}^{n} a_{mij}\ u_i\right)$关于$u$是严格凸函数．又因为$c_{mj}>0$，所以 $c_{mj}\ \exp\left(\sum_{i=1}^{n} a_{mij}\ u_i\right)$关于$u$也是凸函数，因而$Y_m(u)$也是凸函数．由此可知$1-Y_m(u)$是凹函数．因而($CP_2$)是凸规划．

凸规划(CP_2)的任一局部最优解都是它的总体最优解，而(GP)与(CP_2)是等价的，因而(GP)的任一局部最优解都是(GP)的总体最优解．

§4.3　凸规划最优解的充分必要条件

在本节中研究凸规划的最优解的充分必要条件．

定理 4.3.1　设$S\subset\mathscr{R}^n$为非空凸集，$f:S\to\mathscr{R}$为凸函数，则$x^{(0)}\in S$为凸规划（4.1-1）最优解的充分必要条件是：在点$x^{(0)}$存在一个次梯度ξ使得

$$\xi^T(x-x^{(0)})\geqslant 0,\quad \forall x\in S. \tag{4.3-1}$$

证　必要性．设$x^{(0)}\in S$是最优解．于是$\forall x\in S$均有

$$f(x)\geqslant f(x^{(0)})=f(x^{(0)})+\xi^T(x-x^{(0)}).$$

其中$\xi=0\in\mathscr{R}^n$. 因而$\xi=0$是$f(\boldsymbol{x})$在$\boldsymbol{x}^{(0)}$的一个次梯度，而且满足式(4.3-1).

充分性. 设在$\boldsymbol{x}^{(0)}$存在一个次梯度ξ使不等式(4.3-1)成立. 再根据次梯度的定义得

$$
\begin{aligned}
f(\boldsymbol{x}) &\geqslant f(\boldsymbol{x}^{(0)})+\xi^T(\boldsymbol{x}-\boldsymbol{x}^{(0)}) \\
&\geqslant f(\boldsymbol{x}^{(0)}), \qquad \forall \boldsymbol{x}\in S,
\end{aligned}
$$

所以$\boldsymbol{x}^{(0)}$为凸规划(4.1-4)的最优解.

定理 4.3.2 设$S\subset\mathscr{R}^n$为非空凸集，$f:S\to\mathscr{R}$为凸函数，而且在S的内点$\boldsymbol{x}^{(0)}$可微，则$\boldsymbol{x}^{(0)}$为凸规划(4.1-1)的最优解的充分必要条件是

$$
\nabla f(\boldsymbol{x}^{(0)})=0. \tag{4.3-2}
$$

证 因为$f(\boldsymbol{x})$是凸函数，而且在$\boldsymbol{x}^{(0)}$可微，所以$\nabla f(\boldsymbol{x}^{(0)})$是$f(\boldsymbol{x})$在$\boldsymbol{x}^{(0)}$的唯一次梯度（根据定理3.7.3）. 若等式（4.3-2)成立，则条件（4.3-1）满足，因而$\boldsymbol{x}^{(0)}$是凸规划（4.1-1）的最优解. 反之，若$\boldsymbol{x}^{(0)}$是最优解，由定理4.3.1知

$$
\nabla f(\boldsymbol{x}^{(0)})^T(\boldsymbol{x}-\boldsymbol{x}^{(0)})\geqslant 0, \quad \forall \boldsymbol{x}\in S.
$$

因为$\boldsymbol{x}^{(0)}$是S的内点，所以必有$\nabla f(\boldsymbol{x}^{(0)})=0$.

下面考虑著名的Kuhn-Tucker最优性必要条件.

定理 4.3.3 (Kuhn-Tucker**最优性必要条件**)[5] 设$f(\boldsymbol{x})$, $g_j(\boldsymbol{x})$, $h_i(\boldsymbol{x})\in C^1$, $j=1,\cdots,p$; $i=1,\cdots,m$; $\boldsymbol{x}^{(0)}$是约束的正则点. 若$\boldsymbol{x}^{(0)}$是(MP)的局部最优解，则存在$\lambda\in\mathscr{R}^m$, $\mu\in\mathscr{R}^p$使

$$
\begin{aligned}
\nabla f(\boldsymbol{x}^{(0)}) &+ \sum_{i=1}^{m}\lambda_i\nabla h_i(\boldsymbol{x}^{(0)}) \\
&- \sum_{j=1}^{p}\mu_j\nabla g_j(\boldsymbol{x}^{(0)})=0,
\end{aligned} \tag{4.3-3}
$$

$$
\mu_j g_j(\boldsymbol{x}^{(0)})=0, \quad j=1,\cdots,p, \tag{4.3-4}
$$

$$\mu_j \geqslant 0, \quad j=1, \cdots, p. \tag{4.3-5}$$

（证明从略）

上述条件称为Kuhn-Tucker最优性必要条件，简称为K-T条件．满足上述条件的点$x^{(0)}$称为K-T点．K-T条件是$x^{(0)}$为局部最优解的必要条件，对于一般的数学规划（MP），它并非充分条件（参看例1与例2）．

若在(MP)中只含等式约束（即$p=0$），则(MP)就是在数学分析中讨论过的条件极值问题．下式

$$L(x, \lambda, \mu)=f(x)+\sum_{i=1}^{m} \lambda_i h_i(x)-\sum_{j=1}^{p} \mu_j g_j(x)$$

称为（广义）Lagrange函数．称$\lambda=(\lambda_1, \lambda_2, \cdots, \lambda_m)^T$与$\mu=(\mu_1, \mu_2, \cdots, \mu_p)^T$为约束(4.1-4)与(4.1-5)的Lagrange乘子（向量）．

在数学分析中用Lagrange乘数法解条件极值问题所得点是K-T点，但不一定是极小点．如果已知问题的极小点存在，而且K-T点又是唯一的，那么K-T点就是极小点．

例1 $\min f=2xy$,

$$\text{s.t.} \quad 3x+2y=6. \tag{4.3-6}$$

解 先用Lagrange乘数法试之．令

$$L=2xy+\lambda(3x+2y-6).$$

依次对x，y，λ求导，并且令它们为零得

$$L_x=2y+3\lambda=0,$$
$$L_y=2x+2\lambda=0,$$
$$L_\lambda=3x+2y-6=0.$$

解之得$x=1$，$y=3/2$，$\lambda=-1$．代入目标函数得$f=3$．为判断是否为最优值，换一种解法试之．由约束(4.3-6)将y解出，代入目标函数得

$$f=2x\left(3-\frac{3}{2}x\right)=3-3(x-1)^2.$$

当$x\to+\infty$（或$-\infty$）时，$f\to-\infty$，因而此题无最优解. 前面用Lagrange乘数法所得的K-T点$(1, 3/2)^T$是极大点，不是极小点.

例2 设有非线性规划问题

$$\min \quad f=-5x_1^2+x_2^2x_3,$$

$$\text{s.t.} \quad g_1=2-5x_1x_2^{-1}+3x_2^{-1}x_3^2\geqslant 0,$$

$$x_1,\ x_3\geqslant 0,\ x_2\geqslant 0.5.$$

试验证点$\boldsymbol{x}^{(0)}=(3, 6, 1)^T$是上述非线性规划的一个K-T点，再证明$\boldsymbol{x}^{(0)}$不是极小点.

证 容易验证$\boldsymbol{x}^{(0)}$是可行点. 在$\boldsymbol{x}^{(0)}$只有第一个约束是起作用的. 因为

$$\nabla g_1(\boldsymbol{x}^{(0)})=(-5/6,\ 1/3,\ 1)^T\neq\boldsymbol{0},$$

所以$\boldsymbol{x}^{(0)}$是正则点.

上述问题的K-T条件为

$$\begin{aligned}
&-10x_1+5\mu_1x_2^{-1}-\mu_2=0,\\
&2x_2x_3-\mu_1(5x_1x_2^{-2}-3x_2^{-2}x_3^2)-\mu_3=0,\\
&x_2^2-\mu_1(6x_2^{-1}x_3)-\mu_4=0,\\
&\mu_1(2-5x_1x_2^{-1}+3x_2^{-1}x_3^2)=0,\\
&\mu_2x_1=0,\\
&\mu_3x_3=0,\\
&\mu_4(x_2-0.5)=0,\\
&\mu_1,\ \cdots,\ \mu_4\geqslant 0.
\end{aligned}$$

令$\boldsymbol{\mu}=(36, 0, 0, 0)^T$，则$\boldsymbol{x}^{(0)}$与$\boldsymbol{\mu}$满足以上条件，所以$\boldsymbol{x}^{(0)}$是一个K-T点. 相应的目标函数值为$f(\boldsymbol{x}^{(0)})=-9$.

容易验证：$\boldsymbol{x}^{(1)}=(6, 6, 3)^T$是一个可行点，相应的目标函数值为$f_1=f(\boldsymbol{x}^{(1)})=-72$. 所以$\boldsymbol{x}^{(0)}$不是总体极小点.

还可以进一步验证$x^{(0)}$不是局部极小点. 令

$$x=(3+6t,\ 6,\ 1+5t)^T,$$

相应地有

$$f(x)=-5(3+6t)^2+36(1+5t)$$
$$=-9-180t^2,$$

$$g_1(x)=2-\frac{5}{6}(3+6t)+\frac{1}{2}(1+5t)^2$$
$$=12.5t^2.$$

当$t>0$时，点x是可行的，而且相应的目标函数值$f(x)<-9$，所以$x^{(0)}$不是局部极小点.

对于凸规划，K-T条件不仅是极小点的必要条件，而且也是充分条件.

定理 4.3.4 设(MP)是凸规划，$f(x)$，$g_j(x)\in C^1$，$j=1$，$\cdots$，p，$x^{(0)}$为可行解. 若存在$\lambda\in\mathscr{R}^m$，$\mu\in\mathscr{R}^p$满足K-T条件（式(4.3-3)～(4.3-5)），则$x^{(0)}$是凸规划（MP）的总体最优解.

证 设$x\in\mathscr{R}^n$为凸规划(MP)的任一可行解，则条件(4.1-4)与(4.1-5)满足. 又因为$\mu_j\geqslant 0$，$j=1$，$\cdots$，p，所以

$$f(x)\geqslant f(x)+\sum_{i=1}^{m}\lambda_i h_i(x)-\sum_{j=1}^{p}\mu_j g_j(x). \tag{4.3-7}$$

因为$f(x)$，$-g_j(x)$为凸函数，根据定理3.2.2得

$$f(x)\geqslant f(x^{(0)})+\nabla f(x^{(0)})^T(x-x^{(0)}), \tag{4.3-8}$$

$$-g_j(x)\geqslant -g_j(x^{(0)})-\nabla g_j(x^{(0)})^T(x-x^{(0)}), \tag{4.3-9}$$

$$j=1,\ \cdots,\ p.$$

对于仿射函数$h_i(x)$有

$$h_i(x)=h_i(x^{(0)})+\nabla h_i(x^{(0)})^T(x-x^{(0)}),\quad i=1,\ \cdots,\ m. \tag{4.3-10}$$

由式（4.3-7）～（4.3-10）得

$$
\begin{aligned}
f(\boldsymbol{x}) \geqslant f(\boldsymbol{x}^{(0)}) &+ \sum_{i=1}^{m} \lambda_i h_i(\boldsymbol{x}^{(0)}) - \sum_{j=1}^{p} \mu_j g_j(\boldsymbol{x}^{(0)}) \\
&+ \Big[\nabla f(\boldsymbol{x}^{(0)}) + \sum_{i=1}^{m} \lambda_i \nabla h_i(\boldsymbol{x}^{(0)}) \\
&- \sum_{j=1}^{p} \mu_j \nabla g_j(\boldsymbol{x}^{(0)})\Big]^T (\boldsymbol{x} - \boldsymbol{x}^{(0)}),
\end{aligned}
$$

因为$\boldsymbol{x}^{(0)}$为可行解，而且$\boldsymbol{x}^{(0)}$与$\boldsymbol{\lambda}$，$\boldsymbol{\mu}$满足K-T条件，所以上式化为

$$f(\boldsymbol{x}) \geqslant f(\boldsymbol{x}^{(0)}).$$

再由$\boldsymbol{x}$的任意性知，$\boldsymbol{x}^{(0)}$为(MP)的总体最优解.

§4.4 共轭函数与凸规划的对偶理论

4.4.1 共轭函数

定义 4.4.1 设$f: E \to \overline{\mathscr{R}}$，则由下式确定的函数$f^*: E' \to \overline{\mathscr{R}}$

$$f^*(\boldsymbol{y}) = \sup_{\boldsymbol{x} \in E} \{(\boldsymbol{x} | \boldsymbol{y}) - f(\boldsymbol{x})\}, \quad \boldsymbol{y} \in E' \tag{4.4-1}$$

称为$f(\boldsymbol{x})$的**共轭函数**（conjugate function）.

定义 4.4.2 设$g: E' \to \overline{\mathscr{R}}$，则由下式确定的函数$g^*: E \to \overline{\mathscr{R}}$，

$$g^*(\boldsymbol{x}) = \sup_{\boldsymbol{y} \in E'} \{(\boldsymbol{x} | \boldsymbol{y}) - g(\boldsymbol{y})\}, \quad \boldsymbol{x} \in E \tag{4.4-2}$$

称为$g(\boldsymbol{y})$的**共轭函数**.

$f(\boldsymbol{x})$的共轭函数的共轭称为$f(\boldsymbol{x})$的**双极**（bipolar）.

若$f^*(\boldsymbol{y})$是有限实值，根据上确界的定义，$f^*(\boldsymbol{y})$等于满足下式的最小实数α：

$$f(\boldsymbol{x}) \geqslant (\boldsymbol{x} | \boldsymbol{y}) - \alpha, \quad \boldsymbol{x} \in E. \tag{4.4-3}$$

4.4.2 Fenchel对偶定理

在凸规划的研究中，有时转化为它的对偶问题处理．构造对偶问题有多种方法，因而对偶定理也有不同形式．如Rockafellar对偶理论[1]，Wolfe对偶理论[12]，Fenchel对偶定理[2]．首先介绍后者，然后讨论Wolfe对偶理论.

定理 4.4.1 （Fenchel对偶定理） 设$f:E\to\mathscr{R}\cup\{+\infty\}$为正常凸函数，$g:E\to\mathscr{R}\cup\{+\infty\}$为正常凹函数．若

1° $f(\boldsymbol{x})$或$g(\boldsymbol{x})$在某一点$\boldsymbol{x}^{(0)}\in\mathrm{dom}(f)\cap\mathrm{dom}(g)$连续；或

2° $E=\mathscr{R}^n$且$\mathrm{ri}(\mathrm{dom}(f))\cap\mathrm{ri}(\mathrm{dom}(g))\neq\Phi$,

则

$$\inf_{x\in E}\{f(\boldsymbol{x})-g(\boldsymbol{x})\}=\max_{y\in E'}\{g^*(\boldsymbol{y})-f^*(\boldsymbol{y})\}. \qquad (4.4\text{-}4)$$

（证明从略）

4.4.3 Wolfe对偶理论

Wolfe对偶理论较简明，便于应用，现予以较详细地讨论．考虑以下凸规划

(CP) $\min f(\boldsymbol{x}),\ \mathrm{s.t.}\,g_j(\boldsymbol{x})\geqslant 0,\ j=1,\cdots,p,$

其中$f(\boldsymbol{x})$为凸函数，$g_j(\boldsymbol{x})$为凹函数，$f(\boldsymbol{x})$，$g_j(\boldsymbol{x})\in C^1$，$\boldsymbol{x}\in\mathscr{R}^n$

定义 4.4.3 设(CP)的Lagrange函数为

$$L(\boldsymbol{x},\ \lambda)=f(\boldsymbol{x})-\sum_{j=1}^{p}\lambda_j g_j(\boldsymbol{x}).$$

则称以下非线性规划

(DP) $\max L(\boldsymbol{x},\ \lambda),\ \mathrm{s.t.}\,\nabla_x L(\boldsymbol{x},\ \lambda)=0,\ \lambda\geqslant 0$

为凸规划(CP)的**对偶规划**，称(CP)为**原始规划**.

作为一个特例，考虑以下凸二次规划

$$\min f(\boldsymbol{x})=\frac{1}{2}\boldsymbol{x}^T G\boldsymbol{x}+\boldsymbol{c}^T\boldsymbol{x},\ \mathrm{s.t.}\ A\boldsymbol{x}\geqslant\boldsymbol{b}.$$

其中G为$n\times n$对称半正定矩阵，A为$m\times n$矩阵，$c\in\mathscr{R}^n$，$b\in\mathscr{R}^m$.
其对偶规划为

$$\max L(x,\ \lambda)=\frac{1}{2}x^TGx+c^Tx-\lambda^T(Ax-b),$$

$$\text{s.t.}\ \nabla_x L=Gx+c-A^T\lambda=0,\ \lambda\geqslant 0.$$

因为

$$L(x,\lambda)=-\frac{1}{2}x^TGx+x^TGx+c^Tx-(Ax)^T\lambda+b^T\lambda$$

$$=-\frac{1}{2}x^TGx+(Gx+c-A^T\lambda)^Tx+b^T\lambda$$

$$=-\frac{1}{2}x^TGx+b^T\lambda,$$

所以上述对偶规划简化为

$$\max L(x,\ \lambda)=-\frac{1}{2}x^TGx+b^T\lambda,$$

$$\text{s.t.}\quad Gx+c-A^T\lambda=0,\ \lambda\geqslant 0.$$

定理 4.4.2(**弱对偶**) 设x为原始问题(CP)的所有可行解，x，λ为对偶问题(DP)的所有可行解. 则

$$\inf f(x)\geqslant\sup L(x,\ \lambda).$$

证 设y为原始问题(CP)的可行解，x，λ为对偶问题(DP)的可行解. 由$f(x)$的凸性得

$$f(y)\geqslant f(x)+\nabla f(x)^T(y-x).$$

因为x，λ为对偶问题(DP)的可行解，所以

$$\nabla_x L=\nabla f(x)-\sum_{j=1}^{p}\lambda_j\nabla g_j(x)=0,\ \text{且}\ \lambda_j\geqslant 0,\ j=1,\cdots,p$$

于是

$$f(y)\geqslant f(x)+\sum_{j=1}^{p}\lambda_j\nabla g_j(x)^T(y-x)$$

$$\geqslant f(\boldsymbol{x})+\sum_{j=1}^{p}\lambda_j(g_j(\boldsymbol{y})-g_j(\boldsymbol{x}))\text{（因为}g_j(\boldsymbol{x})\text{为凹函数）}$$

$$\geqslant f(\boldsymbol{x})-\sum_{j=1}^{p}\lambda_j g_j(\boldsymbol{x})\quad\text{（因为}g_j(\boldsymbol{y})\geqslant 0\text{）}$$

$$=L(\boldsymbol{x},\ \boldsymbol{\lambda}),$$

因而

$$\inf f(\boldsymbol{y})\geqslant \sup(L(\boldsymbol{x},\ \boldsymbol{\lambda}),\ \text{即}$$

$$\inf f(\boldsymbol{x})\geqslant \sup L(\boldsymbol{x},\ \boldsymbol{\lambda}).$$

注 由此定理可知：若原始问题无界，即目标函数 $f(\boldsymbol{x})$无下界，则对偶问题不可行；若对偶问题无界，即$L(\boldsymbol{x},\ \boldsymbol{\lambda})$无上界，则原始问题不可行.

定理 4.4.3（Wolfe对偶定理） 设$\boldsymbol{x}^*$为原始问题(CP)的最优解. 若$\boldsymbol{x}^*$为约束的正则点，则存在$\boldsymbol{\lambda}^*\geqslant 0$使$\boldsymbol{x}^*$，$\boldsymbol{\lambda}^*$为对偶问题(DP)的最优解，且$L(\boldsymbol{x}^*,\ \boldsymbol{\lambda}^*)=f(\boldsymbol{x}^*)$.

证 因为$\boldsymbol{x}^*$是原始问题(CP)的最优解，且为约束的正则点，所以存在$\boldsymbol{\lambda}^*\geqslant 0$满足K-T条件

$$\nabla_x L(\boldsymbol{x}^*,\ \boldsymbol{\lambda}^*)=\nabla f(\boldsymbol{x}^*)-\sum_{j=1}^{p}\lambda_j^*\nabla g_j(\boldsymbol{x}^*)=0,$$

$$\lambda_j^* g_j(\boldsymbol{x}^*)=0,\ j=1,\ \cdots,\ p.$$

因而$\boldsymbol{x}^*$，$\boldsymbol{\lambda}^*$是(DP)的可行解，且

$$L(\boldsymbol{x}^*,\ \boldsymbol{\lambda}^*)=f(\boldsymbol{x}^*)-\sum_{j=1}^{p}\lambda_j^* g_j(\boldsymbol{x}^*)=f(\boldsymbol{x}^*).$$

再设$\boldsymbol{x}$，$\boldsymbol{\lambda}$为(DP)的任一可行解. 由定理4.4.2可得

$$L(\boldsymbol{x},\ \boldsymbol{\lambda})\leqslant f(\boldsymbol{x}^*)=L(\boldsymbol{x}^*,\ \boldsymbol{\lambda}^*),$$

所以$\boldsymbol{x}^*$，$\boldsymbol{\lambda}^*$为(DP)的最优解.

（关于含等式约束的凸规划的对偶规划请读者进一步讨论）

习　　题

1. 设 $h(\boldsymbol{x})=x_1^2-x_2=0$,

$g_1(\boldsymbol{x})=10x_2-x_1^2-x_2^2\geqslant 0$,

$g_2(\boldsymbol{x})=5-x_2\geqslant 0$.

试求出$g_1(\boldsymbol{x})=0$与$g_2(\boldsymbol{x})=0$的交点，并指出其中哪些是正则点.

2. 试判断下列各数学规划是否为凸规划？若是，再求出它的最优解，并计算最优值.

1° $\min \quad f(\boldsymbol{x})=3x_1^2-5x_1x_2+4x_2^2+15x_1-24x_2$

$\text{s.t.} \quad x_1^2+x_2^2\leqslant 16,\ \boldsymbol{x}=(x_1,\ x_2)^T$.

2° $\min \quad f(\boldsymbol{x})=x_1^3+x_2^3+3x_1^2-3x_2^2+4x_3^2-9x_1-8x_3$

$\text{s.t.} \quad \boldsymbol{x}=(x_1,\ x_2,\ x_3)^T\geqslant \boldsymbol{0}$.

3. 试构造一例：它的K-T点不是最优解.

第五章　凸二次规划

凸二次规划(convex quadratic programming)是最重要最基本的一类非线性规划. 本章研究三种常用解法，还研究凸二次规划的灵敏度分析与对偶规划、对偶理论，也要研究与凸二次规划密切相关的线性互补问题.

§5.1　可行点有效集算法

5.1.1　算法的理论基础

本节研究下面的凸二次规划

$$\min \quad f(\boldsymbol{x})=\frac{1}{2}\boldsymbol{x}^T G\boldsymbol{x}+\boldsymbol{c}^T\boldsymbol{x}, \tag{5.1-1}$$

(CQP)

$$\text{s.t.} \quad \boldsymbol{a}_i^T\boldsymbol{x}=b_i, \ i\in I'. \tag{5.1-2}$$

$$\boldsymbol{a}_i^T\boldsymbol{x}\geqslant b_i, \ i\in I. \tag{5.1-3}$$

其中Hesse矩阵G为$n\times n$对称正定矩阵，$\boldsymbol{c}$，a_i，$\boldsymbol{x}\in\mathscr{R}^n$.

因为G是正定矩阵，所以(CQP)为严格凸规划. 可行点$\boldsymbol{x}^{(k)}$是它的最优解的充分必要条件为$\boldsymbol{x}^{(k)}$为K-T点，即在点$\boldsymbol{x}^{(k)}$K-T条件成立.

设S为(CQP)的可行域，$\boldsymbol{x}^{(k)}\in S$，$I_k$为在$\boldsymbol{x}^{(k)}$的有效约束下标集，即$I_k=\{i|\boldsymbol{a}_i^T\boldsymbol{x}^{(k)}=b_i\}$，$A^{(k)}$为以$\boldsymbol{a}_i^T$，$i\in I_k$，为行形成的矩阵，而且是满行秩的，即$\boldsymbol{x}^{(k)}$为正则点，$\boldsymbol{b}^{(k)}$为由相应的$b_i$组成的列向量，$\boldsymbol{g}^{(k)}=\nabla f(\boldsymbol{x}^{(k)})=G\boldsymbol{x}^{(k)}+\boldsymbol{c}$.

$I'\subset I_k$. 为便于讨论，不妨设$I'=\Phi$，即凸二次规划(CQP)只含不等式约束(5.1-3).

若$\tilde{x}=-G^{-1}c\in S$，则$\tilde{x}$就是(CQP)的唯一最优解，若$\tilde{x}\overline{\in}S$，则下述定理成立.

定理 5.1.1 设$\tilde{x}\overline{\in}S$，则(CQP)的最优解x^*必位于S的边界上.

证 设x为S的任一内点，y为以x、$\tilde{x}$为端点的线段与S的边界的交点，则存在$\alpha\in(0,1)$使$y=\alpha x+(1-\alpha)\tilde{x}$. 由$f(x)$的凸性得

$$f(y)\leqslant\alpha f(x)+(1-\alpha)f(\tilde{x})<f(x),$$

所以内点x不是(CQP)的最优解. 因而(CQP)的最优解x^*必位于可行域S的边界上.

定理 5.1.2 设$x^{(k)}\in S$. 若$x^{(k)}$为凸二次规划

$$(\mathrm{P}_1)\qquad \min\quad f(x)=\frac{1}{2}x^TGx+c^Tx,$$

$$\text{s.t.}\quad A^{(k)}x=b^{(k)}$$

的最优解，且所有Lagrange乘子$\lambda_i\geqslant0$，则$x^{(k)}$为凸二次规划(CQP)的最优解.

证 由假设可得

$$g^{(k)}-\sum_{i\in I_k}\lambda_i a_i=0,\ \lambda_i\geqslant0,\ \forall i\in I_k.$$

令$\lambda_i=0$，$i\in I\setminus I_k$，于是得

$$g^{(k)}-\sum_{i\in I}\lambda_i a_i=0.$$

$$\lambda_i(a_i^Tx^{(k)}-b_i)=0,\ i\in I,$$

$$\lambda_i\geqslant0\qquad\quad ,\ i\in I,$$

即K-T条件成立，因而$x^{(k)}$是(CQP)的最优解.

易知，若$x^{(k)}$是(CQP)的最优解，则$x^{(k)}$也是(P_1)的最优

解.

定理 5.1.3 设$\boldsymbol{x}^{(k)}\in S$，$\boldsymbol{x}=\boldsymbol{x}^{(k)}+\boldsymbol{d}$，则$\boldsymbol{x}^{(k)}$是$(\mathrm{P}_1)$的最优解等价于$\boldsymbol{d}^{(k)}=0$是凸二次规划

$$(\mathrm{P}_2)\qquad \min\ \varphi(\boldsymbol{d})=\frac{1}{2}\boldsymbol{d}^T G\boldsymbol{d}+\boldsymbol{g}^{(k)T}\boldsymbol{d},$$

$$\text{s.t.}\ A^{(k)}\boldsymbol{d}=0$$

的最优解.

证 因为

$$\begin{aligned}
f(\boldsymbol{x})&=\frac{1}{2}(\boldsymbol{x}^{(k)}+\boldsymbol{d})^T G(\boldsymbol{x}^{(k)}+\boldsymbol{d})+\boldsymbol{c}^T(\boldsymbol{x}^{(k)}+\boldsymbol{d})\\
&=\frac{1}{2}\boldsymbol{d}^T G\boldsymbol{d}+(G\boldsymbol{x}^{(k)})^T\boldsymbol{d}+\boldsymbol{c}^T\boldsymbol{d}+\frac{1}{2}\boldsymbol{x}^{(k)T}G\boldsymbol{x}^{(k)}+\boldsymbol{c}^T\boldsymbol{x}^{(k)}\\
&=\frac{1}{2}\boldsymbol{d}^T G\boldsymbol{d}+\boldsymbol{g}^{(k)T}\boldsymbol{d}+f(\boldsymbol{x}^{(k)})\\
&=\varphi(\boldsymbol{d})+f(\boldsymbol{x}^{(k)}),
\end{aligned}\tag{5.1-4}$$

而且$A^{(k)}\boldsymbol{x}=\boldsymbol{b}^{(k)}$等价于$A^{(k)}\boldsymbol{d}=0$（因为$\boldsymbol{x}=\boldsymbol{x}^{(k)}+\boldsymbol{d}$且$A^{(k)}\boldsymbol{x}^{(k)}=\boldsymbol{b}^{(k)}$），所以$\boldsymbol{x}^{(k)}$是$(\mathrm{P}_1)$的最优解等价于$\boldsymbol{d}^{(k)}=0$是$(\mathrm{P}_2)$的最优解.

根据定理5.1.2与定理5.1.3，为求解（CQP），可先取可行点$\boldsymbol{x}^{(k)}$，形成(P_2)并解之. 若(P_2)的最优解$\boldsymbol{d}^{(k)}=0$，而且Lagrange乘子λ_i均非负，则$\boldsymbol{x}^{(k)}$是(CQP)的最优解.

下面再进一步研究如下两个问题.

1. 若(P_2)的最优解$\boldsymbol{d}^{(k)}\neq 0$（此时$\varphi(\boldsymbol{d}^{(k)})<0$），怎么办?

2. 若$\boldsymbol{d}^{(k)}=0$，而Lagrange乘子λ_i中有负的，怎么办?

问题1分两种情形:

A. $\boldsymbol{x}^{(k)}+\boldsymbol{d}^{(k)}\in S$

在这种情形下可以证明以下定理.

定理 5.1.4 设 $x^{(k)}+d^{(k)}\in S$. 若 $x^{(k+1)}=x^{(k)}+d^{(k)}$, $I_{k+1}=I_k$, $A^{(k+1)}=A^{(k)}$, 则

1° $f(x^{(k+1)})<f(x^{(k)})$;

2° 在(P_2)中将k换成$k+1$,所得新问题(P_2)的最优解 $d^{(k+1)}=0$.

证 因为$d^{(k)}\neq 0$, 此时$\varphi(d^{(k)})<0$, 由式(5.1-4)得

$$f(x^{(k+1)})=f(x^{(k)})+\varphi(d^{(k)})<f(x^{(k)}). \qquad (5.1\text{-}5)$$

因为$d^{(k)}$是(P_2)的最优解, 所以存在λ使

$$A^{(k)T}\lambda=Gd^{(k)}+g^{(k)}=g^{(k+1)},$$

即

$$A^{(k+1)}\lambda=G0+g^{(k+1)},$$

所以$d^{(k+1)}=0$是新问题(P_2)的最优解.

B. $x^{(k)}+d^{(k)}\overline{\in} S$

在这种情形下选$\alpha\in(0, 1)$使$x^{(k+1)}=x^{(k)}+\alpha d^{(k)}\in S$. 因为$x^{(k)}\in S$, $A^{(k)}d^{(k)}=0$, 所以只需

$$a_i^T(x^{(k)}+\alpha d^{(k)})\geqslant b_i, \quad i\overline{\in} I_k. \qquad (5.1\text{-}6)$$

因为$x^{(k)}\in S$, 若$a_i^Td^{(k)}\geqslant 0$, 则$\forall\alpha>0$上式都成立.因而使式(5-6)成立的$\alpha>0$的上界为

$$\bar{\alpha}=\min_{i\overline{\in} I_k}\left\{\frac{a_i^Tx^{(k)}-b_i}{-a_i^Td^{(k)}}\,\middle|\,a_i^Td^{(k)}<0\right\}$$

$$=(a_l^Tx^{(k)}-b_l)/(-a_l^Td^{(k)}). \qquad (5.1\text{-}7)$$

容易证明: 由上式确定的$\bar{\alpha}\in(0, 1)$. 令

$$x^{(k+1)}=x^{(k)}+\bar{\alpha}d^{(k)}$$

$$=\bar{\alpha}(x^{(k)}+d^{(k)})+(1-\bar{\alpha})x^{(k)}. \qquad (5.1\text{-}8)$$

由$f(x)$的凸性得

$$f(\boldsymbol{x}^{(k+1)}) \leqslant \bar{\alpha} f(\boldsymbol{x}^{k}+\boldsymbol{d}^{(k)})+(1-\bar{\alpha}) f(\boldsymbol{x}^{k}).$$

因为

$$f(\boldsymbol{x}^{(k)}+\boldsymbol{d}^{(k)})=\varphi(\boldsymbol{d}^{(k)})+f(\boldsymbol{x}^{(k)}),\ \text{而且}\ \varphi(\boldsymbol{d}^{(k)})<0,$$

所以

$$f(\boldsymbol{x}^{k+1})<f(\boldsymbol{x}^{k}).$$

容易验证

$$\boldsymbol{a}_i^T \boldsymbol{x}^{k+1}=b_i,\quad i \in I_k;$$

$$\boldsymbol{a}_l^T \boldsymbol{x}^{k+1}=b_l.$$

可以将上述讨论概括为以下定理.

定理 5.1.5 设$\boldsymbol{x}^{(k)} \in S$, $\boldsymbol{d}^{(k)} \neq 0$, $\boldsymbol{x}^{(k)}+\boldsymbol{d}^{(k)} \overline{\in} S$, $\bar{\alpha}$由式(5.1-7)确定，$\boldsymbol{x}^{(k+1)}$由式(5.1-8)确定，则

1° $\boldsymbol{x}^{(k+1)} \in S$, $f(\boldsymbol{x}^{k+1})<f(\boldsymbol{x}^{k})$;

2° $\boldsymbol{a}_l^T \boldsymbol{x}^{(k+1)}=b_l$, $\boldsymbol{a}_i^T \boldsymbol{x}^{(k+1)}=b_i$, $i \in I_k$.

问题 2 按以下办法处理：令

$$\lambda_s^{(k)}=\min_{i \in I_k}\{\lambda_i^{(k)}\}. \tag{5.1-9}$$

将s从有效集I_k中去掉（若满足上式的s多于一个，每次只能从I_k中去掉一个元素），从$A^{(k)}$中将相应的行去掉，仍然记为$A^{(k)}$. 而$\boldsymbol{x}^{(k)}$, $\boldsymbol{g}^{(k)}$不变. 形成新的凸二次规划(P_2)再解之. 可以证明：新问题(P_2)的最优解$\boldsymbol{d}^{(k)}$必不为0.

5.1.2 计算步骤

设在(CQP)中$I'=\Phi$, $\boldsymbol{x}^{(0)}$为初始可行点.

1° 令$k=0$,确定有效集$I_k=\{i \mid \boldsymbol{a}_i^T \boldsymbol{x}^{(k)}=b_i\}$，形成矩阵$A^{(k)}$;

2° 计算梯度$\boldsymbol{g}^{(k)}=\nabla f(\boldsymbol{x}^{(k)})=G\boldsymbol{x}^{(k)}+\boldsymbol{c}$;

3° 解等式约束问题(P_2). 若最优解$\boldsymbol{d}^{(k)} \neq 0$, 转 5°;

4° 解方程组$A^{(k)T}\lambda=\boldsymbol{g}^{(k)}$得 Lagrange 乘子$\lambda^{(k)}$, 并按式(5.1-9)计算$\lambda_s^{(k)}$. 若$\lambda_s^{(k)} \geqslant 0$, 则$\boldsymbol{x}^{(k)}$为(CQP)的最优解$\boldsymbol{x}^*$, 停；否则，令$I_k=I_k \setminus \{s\}$, 从$A^{(k)}$中将$\boldsymbol{a}_s^T$删除，仍然记为

$A^{(k)}$，转3°；

5° 计算α_k与$\boldsymbol{x}^{(k+1)}$.

$$\alpha_k=\min\{1, (\boldsymbol{a}_l^T\boldsymbol{x}^{(k)}-b_l)/(-\boldsymbol{a}_l^T\boldsymbol{d}^{(k)})\},$$

其中 $(\boldsymbol{a}_l^T\boldsymbol{x}^{(k)}-b_l)/(-\boldsymbol{a}_l^T\boldsymbol{d}^{k})$ 按式(5.1-7)算得，

$$\boldsymbol{x}^{k+1}=\boldsymbol{x}^{k}+\alpha_k\boldsymbol{d}^{(k)}.$$

6° 若$\alpha_k<1$，令$I_{k+1}=I_k\cup\{l\}$，将$\boldsymbol{a}_l^T$并入$A^{(k)}$得$A^{(k+1)}$；

7° 令$k=k+1$，转2°.

若$I'\neq\Phi$，等式约束应看作有效约束，$I'\subset I_k$，但是，计算$^{(k)}$时不考虑这些约束.

5.1.3 收敛性

定理 5.1.6 若G为对称正定矩阵，而且在每次迭代中所得矩阵$A^{(k)}$都是满行秩的，则可行点有效集算法在有限步内终止于(CQP)的最优解$\boldsymbol{x}^*$.

证 若$\boldsymbol{d}^{(k)}\neq 0$，且$0<\alpha_k<1$，则有效集$I_k$改变，而且迭代后目标函数值有所下降. 若$\alpha_k=1$，有效集不变，但是，迭代后得$\boldsymbol{d}^{(k+1)}=0$；若$\lambda_i^{(k)}\geqslant 0$，则迭代终止，得(CQP)的最优解$\boldsymbol{x}^*$；若$\lambda_i^{(k)}<0$，有效集$I_k$改变，但是再迭代一次所得$\boldsymbol{d}^{(k)}\neq 0$. 令$\boldsymbol{x}^{(k+1)}=\boldsymbol{x}^{(k)}+\bar{\alpha}\boldsymbol{d}^{(k)}$，则目标函数值有所下降，$f(\boldsymbol{x}^{(k+1)})<f(\boldsymbol{x}^{(k)})$. 所以，有效集$I_k$不能连续重复. 由$G$的正定性知，(CQP)的最优解存在且唯一. 又因为(CQP)的有效集I_k个数有限（不超过$\sum_{k=1}^{m}C_m^k$，若$I=\{1,\cdots,m\}$），所以，经过有限次迭代必可得到(CQP)的最优解$\boldsymbol{x}^*$.

5.1.4 关于求解等式约束问题(P_2)的消元法

关于等式约束问题(P_2)常用消元法求解. 若$A^{(k)}$的行数$m=n$，则(P_2)的最优解$\boldsymbol{d}^{(k)}=0$，所以不妨设$m<n$. 将$A^{(k)}$分为两块：$A^{(k)}=(B, N)$，而$|B|\neq 0$. 将$\boldsymbol{g}^{(k)}$，$\boldsymbol{d}$，G也作相应的

划分:

$$\boldsymbol{g}^k=\begin{pmatrix}\boldsymbol{g}_B\\ \boldsymbol{g}_N\end{pmatrix},\quad \boldsymbol{d}=\begin{pmatrix}\boldsymbol{d}_B\\ \boldsymbol{d}_N\end{pmatrix},\quad G=\begin{pmatrix}G_B & G_{12}\\ G_{21} & G_N\end{pmatrix},$$

其中$\boldsymbol{g}_B$, $\boldsymbol{d}_B\in\mathscr{R}^m$, $\boldsymbol{g}_N$, $\boldsymbol{d}_N\in\mathscr{R}^{n-m}$, G_B为$m\times m$对称正定矩阵,G_N为$(n-m)\times(n-m)$对称矩阵,$G_{21}=G_{12}^T$为$(n-m)\times m$矩阵.

可以将约束$A^{(k)}\boldsymbol{d}=0$改写为:$B\boldsymbol{d}_B+N\boldsymbol{d}_N=0$. 由此得

$$\boldsymbol{d}_B=-B^{-1}N\boldsymbol{d}_N. \tag{5.1-10}$$

将它代入目标函数$\varphi(\boldsymbol{d})$中,于是(P_2)化为关于$\boldsymbol{d}_N$的无约束问题

$$\min\ \psi(\boldsymbol{d}_N)=\frac{1}{2}\boldsymbol{d}_N^T\widehat{G}\boldsymbol{d}_N+\widehat{\boldsymbol{g}}^T\boldsymbol{d}_N,$$

其中

$$\widehat{G}=G_N+(B^{-1}N)^TG_BB^{-1}N-2G_{21}B^{-1}N, \tag{5.1-11}$$

$$\widehat{\boldsymbol{g}}=\boldsymbol{g}_N-(B^{-1}N)^T\boldsymbol{g}_B. \tag{5.1-12}$$

若$\widehat{G}$是正定矩阵,则$\psi(\boldsymbol{d}_N)$有唯一极小点$\boldsymbol{d}^*_N=-\widehat{G}^{-1}\widehat{\boldsymbol{g}}$. 将它代入式(5.1-10)得$\boldsymbol{d}^*_B=B^{-1}N\widehat{G}^{-1}\widehat{\boldsymbol{g}}$. 于是得($P_2$)的最优解

$$\boldsymbol{d}^k=\begin{pmatrix}B^{-1}N\ \widehat{G}^{-1}\ \widehat{\boldsymbol{g}}\\ -\widehat{G}^{-1}\widehat{\boldsymbol{g}}\end{pmatrix}. \tag{5.1-13}$$

5.1.5 举例

例 $\min\ f(\boldsymbol{x})=\frac{1}{2}x_1^2+\frac{1}{2}x_2^2-x_1-2x_2,$

$$\text{s.t.}\quad 2x_1+3x_2\leqslant 6,\qquad x_1+4x_2\leqslant 5,$$
$$x_1\geqslant 0,\qquad x_2\geqslant 0.$$

解 用-1乘前两个不等式的两端，化为"$\geqslant$"型的. 目标函数$f(x)$的Hesse矩阵与梯度如下

$$G=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix},\qquad \nabla f(\boldsymbol{x})=(x_1-1,\ x_2-2)^T.$$

1° 取初始可行点$\boldsymbol{x}^{(0)}=(3,0)^T$，于是，$I_0=\{1,4\}$，

$$A^{(0)}=\begin{pmatrix}-2 & -3\\ 0 & 1\end{pmatrix},\quad \boldsymbol{g}^{(0)}=\begin{pmatrix}2\\ -2\end{pmatrix}.$$

等式约束问题（P_2）为

$$\min\ \varphi(\boldsymbol{d})=\frac{1}{2}d_1^2+\frac{1}{2}d_2^2+2d_1-2d_2,$$

$$\text{s.t.}\quad -2d_1-3d_2=0,$$
$$d_2=0.$$

解之得最优解$\boldsymbol{d}^{(0)}=(0,\ 0)^T$. 再解方程组

$$\begin{cases}-2\lambda_1 \qquad\quad =2,\\ -3\lambda_1+\lambda_4=-2\end{cases}$$

得 $\boldsymbol{\lambda}^0=(-1,\ -5)^T$.

2° 令 $I_0=I_0\setminus\{4\}=\{1\}$，则$A^{(0)}=(-2,\ -3)$. 于是（$P_2$）为

$$\min\ \varphi(\boldsymbol{d})=\frac{1}{2}d_1^2+\frac{1}{2}d_2^2+2d_1-2d_2,$$

$$\text{s.t.}\quad -2d_1-3d_2=0.$$

用消元法得最优解$\boldsymbol{d}^{(0)}=(-30/13,\ 20/13)^T$.

再计算α_0与$\boldsymbol{x}^{(1)}$:

$$\alpha_0=\min\left\{1,\ (-3+5)\Big/\frac{50}{13},\ (3-0)\Big/\frac{30}{13}\right\}=13/25.$$

$$x^{(1)}=x^{(0)}+\alpha_0 d^{(0)}=(9/5,\ 4/5)^T.$$

3° 令$I_1=I_0\cup\{2\}=\{1,\ 2\}$，则

$$A^{(1)}=\begin{pmatrix}-2 & -3\\ -1 & -4\end{pmatrix},\quad g^{(1)}=(4/5,\ -6/5)^T.$$

(P_2)为

$$\min\ \varphi(d)=\frac{1}{2}d_1^2+\frac{1}{2}d_2^2+\frac{4}{5}d_1-\frac{6}{5}d_2,$$

$$\text{s.t.}\quad -2d_1-3d_2=0,$$
$$-d_1-4d_2=0.$$

解之得最优解 $d^{(1)}=(0,\ 0)^T$.再解方程组

$$\begin{cases}-2\lambda_1-\ \lambda_2=\ \ 4/5,\\ -3\lambda_1-4\lambda_2=-6/5,\end{cases}$$

得 $\lambda^{(1)}=(-22/25,\ 24/25)^T$.

4° 令 $I_1=I_1\setminus\{1\}=\{2\}$，则$A^{(1)}=(-1,\ -4)$，$(P_2)$为

$$\min\ \varphi(d)=\frac{1}{2}d_1^2+\frac{1}{2}d_2^2+\frac{4}{5}d_1-\frac{6}{5}d_2,$$

$$\text{s.t.}\quad -d_1-4d_2=0.$$

解之得最优解 $d^{(1)}=(-88/85,\ 22/85)^T$.再计算$\alpha_1$与$x^{(2)}$:

$$\alpha_1=\min\left\{1,\ \frac{9}{5}\div\frac{88}{85}\right\}=1,$$

$$x^{(2)}=x^{(1)}+\alpha_1 d^{(1)}=(13/17,\ 18/17)^T.$$

5° $g^{(2)}=(-4/17,\ -16/17)^T$，$(P_2)$为

$$\min\ \varphi(d)=\frac{1}{2}d_1^2+\frac{1}{2}d_2^2-(4/17)d_1-(16/17)d_2,$$

$$\text{s.t.}\quad -d_1-4d_2=0.$$

解之得最优解$d^{(2)}=(0, 0)^T$. 再解方程组

$$\begin{cases}-\lambda_1=-4/17,\\-4\lambda_1=-16/17,\end{cases}$$

得$\lambda_1^{(2)}=4/17>0$. 因而$x^{(2)}=(13/17, 18/17)^T$是最优解.

5.1.6 初始可行点$x^{(0)}$的选择

在使用本算法时，初始可行点的选择非常重要，选得好，迭代次数可以大大减少. 对于上一段的例题，若选$x^{(0)}=(1, 1)^T$为初始可行点，则只需迭代一次就可以得到最优解$x^*=(13/17, 18/17)^T$.

关于初始可行点的选择，提出以下两种参考方法.

若$\tilde{x}=-G^{-1}c\in S$，则$\tilde{x}$就是(CQP)的最优解，所以不妨设$\tilde{x}\overline{\in} S$.

方法一 令$\min_i\{a_i^T\tilde{x}-b_i\}=a_l^T\tilde{x}-b_l$，在超平面$a_l^Tx=b_l$上取一个可行点作为初始可行点.

方法二 设$x^{(1)}, x^{(2)}\in S$. 若$\|x^{(1)}-\tilde{x}\|\leqslant\|x^{(2)}-\tilde{x}\|$则以$x^{(1)}$为初始可行点；否则，以$x^{(2)}$为初始可行点.

采用上述方法选择初始可行点，迭代次数往往较少.

对于上一段的例题，$x^{(1)}=(3, 0)^T$与$x^{(2)}=(1, 1)^T$都是可行点，$\tilde{x}=(1, 2)^T$, $\|x^{(1)}-\tilde{x}\|=2\sqrt{2}$, $\|x^{(2)}-\tilde{x}\|=1$. 取$x^{(2)}$作为初始可行点$x^{(0)}$较好.

§5.2 二次规划的K－T条件与线性互补问题

本节研究下面的二次规划

$$\text{(QP)}\quad \begin{aligned}&\min\ f(x)=\frac{1}{2}x^TGx+c^Tx, &(5.2\text{-}1)\\&\text{s.t. } Ax\leqslant b, &(5.2\text{-}2)\\&\qquad x\geqslant 0. &(5.2\text{-}3)\end{aligned}$$

其中$\boldsymbol{c}, \boldsymbol{x} \in \mathscr{R}^n$，$\boldsymbol{b} \in \mathscr{R}^m$，$A$为$m \times n$矩阵，$G$为$n \times n$对称矩阵。

令$\boldsymbol{y}$为关于约束（5.2-2）的松弛变量，$\boldsymbol{u}$，$\boldsymbol{v}$为式（5.2-2）与式（5.2-3）的Lagrange乘子向量，于是二次规划（QP）的K-T条件可以写成

$$
\begin{aligned}
&\boldsymbol{y} + A\boldsymbol{x} = \boldsymbol{b}, \\
&\boldsymbol{v} - A^T\boldsymbol{u} - G\boldsymbol{x} = \boldsymbol{c}, \\
&\boldsymbol{u}^T\boldsymbol{y} = 0, \quad \boldsymbol{v}^T\boldsymbol{x} = 0, \\
&\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{u}, \boldsymbol{v} \geqslant 0.
\end{aligned} \tag{5.2-4}
$$

令

$$
M = \begin{pmatrix} O & -A \\ A^T & G \end{pmatrix}, \quad \boldsymbol{q} = \begin{pmatrix} \boldsymbol{b} \\ \boldsymbol{c} \end{pmatrix}, \tag{5.2-5}
$$

$$
\boldsymbol{w} = \begin{pmatrix} \boldsymbol{y} \\ \boldsymbol{v} \end{pmatrix}, \quad \boldsymbol{z} = \begin{pmatrix} \boldsymbol{u} \\ \boldsymbol{x} \end{pmatrix};
$$

则上述K-T条件又可以写成

(LCP)

$$\boldsymbol{w} - M\boldsymbol{z} = \boldsymbol{q}, \tag{5.2-6}$$

$$\boldsymbol{w}, \boldsymbol{z} \geqslant 0, \tag{5.2-7}$$

$$\boldsymbol{w}^T\boldsymbol{z} = 0. \tag{5.2-8}$$

定义 5.2.1 设M为$p \times p$矩阵，$\boldsymbol{q} \in \mathscr{R}^p$。求$\boldsymbol{w}$，$\boldsymbol{z} \in \mathscr{R}^p$满足式（5.2-6）～（5.2-8），这类问题称为**线性互补问题**（linear complementary problem）。w_j与z_j，$j = 1, \cdots, p$，称为互补变量；式（5.2-8）称为互补条件。

定义 5.2.2 若（$\boldsymbol{w}$，$\boldsymbol{z}$）① 是式(5.2-6)与式（5.2-7）的基本可行解，并且w_j与z_j（$j = 1, \cdots, p$）中恰好有一个为基变量，则称（$\boldsymbol{w}$，$\boldsymbol{z}$）为线性互补问题（LCP）的一个**互补基本可行解**（complementary basic feasible solution）。

设（LCP）是与二次规划（QP）对应的线性互补问题。若（$\boldsymbol{w}$，

①在本章中（w，z）表示一对向量（点）；w，z

$\boldsymbol{z}$）是（LCP）的互补基本可行解，则令$x_i^*=z_{m+j}$，$j=1,\cdots,n$，就得到（QP）的一个K－T点$\boldsymbol{x}^*=(x_1^*, x_2^*, \cdots, x_n^*)^T$.

若G为半正定矩阵，则（QP）为凸二次规划．对于凸规划，K－T条件是最优解的充要条件，K－T点就是最优解．所以凸二次规划（QP）的求解可以转化为与它对应的线性互补问题（LCP）的求解．

若$\boldsymbol{q}\geqslant 0$，则$\boldsymbol{w}=\boldsymbol{q}$，$\boldsymbol{z}=0$就是（LCP）的一个互补基本可行解．否则，引入一个列向量$\boldsymbol{1}$（它的各分量均为1）和一个人工变量z_0将原问题化为具有$2p+1$个变量的新问题，即求$\boldsymbol{w}$，$\boldsymbol{z}$，z_0满足以下各式

$$\boldsymbol{w}-M\boldsymbol{z}-\boldsymbol{1}z_0=\boldsymbol{q}, \tag{5.2-9}$$

$$(\mathrm{P}_3)\qquad \boldsymbol{w}\geqslant 0,\ \boldsymbol{z}\geqslant 0,\ z_0\geqslant 0, \tag{5.2-10}$$

$$\boldsymbol{w}^T\boldsymbol{z}=0. \tag{5.2-11}$$

定义 5.2.3 设（$\boldsymbol{w}$，$\boldsymbol{z}$，z_0）是（P_3）的可行解．若

1° （$\boldsymbol{w}$，$\boldsymbol{z}$，z_0）是式（5.2-9）与式（5.2-10）的一个基本可行解；

2° 对于某个s（$s\in\{1,2,\cdots,p\}$），w_s与z_s都不是基变量；

3° z_0是基变量，而w_j，z_j中恰好有一个是基变量，$j\neq s$，

则称（$\boldsymbol{w}$，$\boldsymbol{z}$，z_0）为（P_3）的一个**准互补基本可行解**（almost complementary basic feasible solution）．

定义 5.2.4 设（$\boldsymbol{w}$，$\boldsymbol{z}$，z_0）为（P_3）的一个准互补基本可行解，其中w_s，z_s均为非基变量．用w_s或z_s替换非z_0的一个基变量，所得的准互补基本可行解称为与它相邻的准互补基本可行解．

每一个准互补基本可行解至多有两个相邻的准互补基本可行解．

（P_3）有一个初始准互补基本可行解

$$z_0=-\min\{q_i \mid 1\leqslant i\leqslant p\},\ z=0,\ w=q+1z_0.$$

按照一定规则迭代后得到一个和它相邻的准互补基本可行解．若经过若干次迭代将z_0排出基，化为零，那么就得到（LCP）的一个互补基本可行解．

§5.3 Lemke互补转轴算法

5.3.1 Lemke互补转轴算法的计算步骤

若$q\geqslant 0$，则（w，z）=（q，0）是（LCP）的一个互补基本可行解，否则

1° 将方程组（5.2-9）的增广矩阵列为初始表（仿单纯形表），并以w_1，…，w_p为基变量．令$q_s=\min\{q_i \mid 1\leqslant i\leqslant p\}$，再令$z_0$进基将$w_s$换出基（即按第$s$行$z_0$列转轴），并记$y_s=z_s$．将与$y_s$对应的列记为$d^{(s)}$．

2° 若$d^{(s)}\leqslant 0$，转6°；否则，按最小比规则确定指标r

$$\frac{q_r}{d_r^{(s)}}=\min\left\{\frac{q_i}{d_i^{(s)}}\,\middle|\,d_i^{(s)}>0\right\}. \tag{5.3-1}$$

3° 若第r个基变量为z_0，转5°；否则

4° 按第r行$d^{(s)}$列转轴．若此时离基变量为z_l（或w_l），则记$y_s=w_l$（或z_l），转2°；

5° 按第r行$d^{(s)}$列转轴（令y_s进基，z_0离基），迭代后得一个互补基本可行解，停．

6° （LCP）无互补基本可行解，停．

当$d^{(s)}\leqslant 0$时，根据当前的表可以确定一个准互补基本可行解（w，z，z_0），还可以确定由式（5.2-9）与式（5.2-10）所定的凸集的极方向d：当前的基变量取$-d^{(s)}$中对应的分量，非基变量y_s取为1，其余非基变量均取零值．射线$\widetilde{R}=\{(w,\ z,\ z_0)+\lambda d^T \mid \lambda\geqslant 0\}$上的点都是（$P_3$）的准互补基本可行解（因而可以将

这种情形简称为“射线终止”）.

5.3.2 算法的理论基础

定理 5.3.1 设(P_3)的每个准互补基本可行解均非退化，即基变量都是正的，那么由Lemke互补转轴算法所产生的点（准互补基本可行解）都不重复，因而Lemke互补转轴算法在有限步内终止.

证 设($\boldsymbol{w}, \boldsymbol{z}, z_0$)是($P_3$)的一个准互补基本可行解，而$w_s$, z_s都是非基变量. 至多有两个和它相邻的准互补基本可行解. 根据非退化假设，它们都不同于($\boldsymbol{w}, \boldsymbol{z}, z_0$). 假设$(\boldsymbol{w}, \boldsymbol{z}, z_0)^{(k+\sigma)}$是第一次重复出现的点

$$(\boldsymbol{w}, \boldsymbol{z}, z_0)^{(k+\sigma)}=(\boldsymbol{w}, \boldsymbol{z}, z_0)^{(k)}.$$

因为$(\boldsymbol{w}, \boldsymbol{z}, z_0)^{(k+\sigma-1)}$与$(\boldsymbol{w}, \boldsymbol{z}, z_0)^{(k+\sigma)}$相邻，所以它也与$(\boldsymbol{w}, \boldsymbol{z}, z_0)^{(k)}$相邻. 和$(\boldsymbol{w}, \boldsymbol{z}, z_0)^{(k)}$相邻的只可能是$(\boldsymbol{w}, \boldsymbol{z}, z_0)^{(k-1)}$与$(\boldsymbol{w}, \boldsymbol{z}, z_0)^{(k+1)}$，所以$(\boldsymbol{w}, \boldsymbol{z}, z_0)^{(k+\sigma-1)}$必等于$(\boldsymbol{w}, \boldsymbol{z}, z_0)^{(k-1)}$或$(\boldsymbol{w}, \boldsymbol{z}, z_0)^{(k+1)}$. 这与$(\boldsymbol{w}, \boldsymbol{z})^{(k+\sigma}$是第一次重复出现的点矛盾. 因而，用Lemke互补转轴算法所产生的点不重复. 另一方面，(P_3)至多具有有限多个(至多为C_{2p+1}^{p}个)准互补基本可行解，所以算法必定在有限步内终止.

定理 5.3.2 假设

1° $\forall \boldsymbol{z}\geqslant\boldsymbol{0}$ 均有 $\boldsymbol{z}^T M\boldsymbol{z}\geqslant 0$，而且 $\boldsymbol{z}^T M\boldsymbol{z}=0$ ($\boldsymbol{z}\geqslant\boldsymbol{0}$) 蕴涵 $(M+M^T)\boldsymbol{z}=\boldsymbol{0}$; (5.3-2)

2° 每一个准互补基本可行解均非退化.

若式(5.2-6)与式(5.2-7)相容，则Lemke互补转轴算法终止于一个互补基本可行解；若不相容，则射线终止[3].

（证明从略）

定理 5.3.3 设G为半正定矩阵，(LCP)为与(QP)对应的线性互补问题.

1° 若($\boldsymbol{w}, \boldsymbol{z}$)为(LCP)的互补基本可行解，则$\boldsymbol{x}^*$ ($x^*=$

z_{m+j}, $j=1$, …, n）是（QP）的最优解：

2° 若（QP）具有最优解，且用Lemke互补转轴算法求解对应的（LCP）时所得每一个准互补基本可行解均非退化，则算法终止于一个互补基本可行解（$\boldsymbol{w}$, $\boldsymbol{z}$），从而可得（QP）的最优解 $\boldsymbol{x}^*$.

证 1° 显然，证明从略.

2° 设 $M=\begin{pmatrix} O & -A \\ A^T & G \end{pmatrix}$, $\boldsymbol{z}=\begin{pmatrix} \boldsymbol{u} \\ \boldsymbol{x} \end{pmatrix}$, $\boldsymbol{u}\in\mathscr{R}^m$, $\boldsymbol{x}\in\mathscr{R}^n$,

则

$$\boldsymbol{z}^T M \boldsymbol{z}=\boldsymbol{x}^T G\boldsymbol{x}\geqslant 0 \quad (\forall \boldsymbol{z}),$$

$$(M+M^T)\boldsymbol{z}=\begin{pmatrix} O & O \\ O & 2G \end{pmatrix}\begin{pmatrix} \boldsymbol{u} \\ \boldsymbol{x} \end{pmatrix}=\begin{pmatrix} \boldsymbol{0} \\ 2G\boldsymbol{x} \end{pmatrix}.$$

设 $\boldsymbol{d}=G\boldsymbol{x}$, $\boldsymbol{z}^T M\boldsymbol{z}=0$, $\lambda>0$. 因为G是半正定矩阵，所以

$$0\leqslant(\boldsymbol{x}-\lambda\boldsymbol{d})^T G(\boldsymbol{x}-\lambda\boldsymbol{d})=\boldsymbol{x}^T G\boldsymbol{x}+\lambda^2\boldsymbol{d}^T G\boldsymbol{d}-2\lambda\|\boldsymbol{d}\|^2$$
$$=\lambda^2\boldsymbol{d}^T G\boldsymbol{d}-2\lambda\|\boldsymbol{d}\|^2 (\text{因为}\boldsymbol{x}^T G\boldsymbol{x}=\boldsymbol{z}^T M\boldsymbol{z}=0),$$

上式两端分别除以λ，然后令 $\lambda\to 0^+$，得$-2\|\boldsymbol{d}\|^2\geqslant 0$. 所以$\boldsymbol{d}=\boldsymbol{0}$，即$G\boldsymbol{x}=\boldsymbol{0}$，因而（$M+M^T$）$\boldsymbol{z}=\boldsymbol{0}$. 条件（5.3-2）满足. 因为（QP）具有最优解，所以式（5.2-6）与式（5.2-7）相容. 再由假设每一个准互补基本可行解均非退化，故根据定理5.3.2，Lemke互补转轴算法终止于一个互补基本可行解，从而得（QP）的最优解$\boldsymbol{x}^*$.

推论 设G为半正定矩阵，若所有准互补基本可行解均非退化，则“射线终止”意味着（QP）无最优解. 或可行域为空集，或可行域无界，目标函数$f(\boldsymbol{x})$无下界.

注 若可行域是非空有界闭集，则（QP）具有最优解.

定理 5.3.4 二次规划

$$\min \quad f(\boldsymbol{x})=\frac{1}{2}\boldsymbol{x}^T G\boldsymbol{x}+\boldsymbol{c}^T\boldsymbol{x},$$

$$\text{s.t.} \quad \boldsymbol{a}_i^T\boldsymbol{x}=b_i, \quad i\in I,$$
$$\boldsymbol{a}_i^T\boldsymbol{x}\leqslant b_i, \quad i\in I'$$

具有最优解的充要条件是：目标函数$f(\boldsymbol{x})$在可行域上有下界[20].

推论 若G为正定矩阵或G为半正定矩阵而且$\boldsymbol{c}=\boldsymbol{0}$，可行域非空，则（QP）具有最优解。

定理 5.3.5 设G无负元素，而且主对角线上的元素均为正，可行域非空．若所有准互补基本可行解均非退化，则算法终止时可得（QP）的一个K－T点[3]（G可以不是半正定矩阵，此时K－T点不一定是最优解）.

5.3.3 举例

例 1 $\min \quad f(\boldsymbol{x})=\frac{1}{2}x_1^2+\frac{1}{2}x_2^2-x_1-2x_2,$

$$\begin{aligned}\text{s.t.} \quad & 2x_1+3x_2\leqslant 6,\\ & x_1+4x_2\leqslant 5,\\ & x_1, x_2 \geqslant 0.\end{aligned}$$

解 已知数据如下

$$G=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}, \quad \boldsymbol{c}=\begin{pmatrix}-1\\ -2\end{pmatrix}, \quad A=\begin{pmatrix}2 & 3\\ 1 & 4\end{pmatrix}, \quad \boldsymbol{b}=\begin{pmatrix}6\\ 5\end{pmatrix}.$$

令

$$M=\begin{pmatrix}O & -A\\ A^T & G\end{pmatrix}, \quad \boldsymbol{q}=\begin{pmatrix}\boldsymbol{b}\\ \boldsymbol{c}\end{pmatrix}, \quad \boldsymbol{w}=\begin{pmatrix}\boldsymbol{y}\\ \boldsymbol{v}\end{pmatrix}, \quad \boldsymbol{z}=\begin{pmatrix}\boldsymbol{u}\\ \boldsymbol{x}\end{pmatrix}.$$

形成线性互补问题（LCP），然后用Lemke互补转轴算法求解（详见表5－1）．经过四次迭代，z_0从基中排出，得一互补基本可行解：

$$\boldsymbol{w}=(22/17, 0, 0, 0)^T,$$
$$\boldsymbol{z}=(0, 4/17, 13/17, 18/17)^T.$$

凸二次规划的最优解为$\boldsymbol{x}^*=(13/17, 18/17)^T$．可以算出最

表 5-1

基	w_1	w_2	w_3	w_4	z_1	z_2	z_3	z_4	z_0	q
w_1	1	0	0	0	0	0	2	3	−1	6
w_2	0	1	0	0	0	0	1	4	−1	5
w_3	0	0	1	0	−2	−1	−1	0	−1	−1
w_4	0	0	0	1	−3	−4	0	−1	(−1)	−2
w_1	1	0	0	−1	3	4	2	4	0	8
w_2	0	1	0	−1	3	4	1	5	0	7
w_3	0	0	1	−1	1	3	−1	(1)	0	1
z_0	0	0	0	−1	3	4	0	1	1	2
w_1	1	0	−4	3	−1	−8	6	0	0	4
w_2	0	1	−5	4	−2	−11	(6)	0	0	2
z_4	0	0	1	−1	1	3	−1	1	0	1
z_0	0	0	−1	0	2	1	1	0	1	1
w_1	1	−1	1	−1	1	3	0	0	0	2
z_3	0	1/6	−5/6	2/3	−1/3	−11/6	1	0	0	1/3
z_4	0	1/6	1/6	−1/3	2/3	7/6	0	1	0	4/3
z_0	0	−1/6	−1/6	−2/3	7/3	(17/6)	0	0	1	2/3
w_1	1	−14/17	20/17	−5/17	−25/17	0	0	0	−18/17	22/14
z_3	0	1/17	−16/17	4/17	20/17	0	1	0	11/17	13/17
z_4	0	4/17	4/17	−1/17	−5/17	0	0	1	−7/17	18/17
z_2	0	−1/17	−1/17	−4/17	14/17	1	0	0	6/17	4/17

优值$f^*=-1173/578=-2.0294$. 因为G是正定矩阵，所以最优解$\boldsymbol{x}^*$是唯一的.

例 2 $\min\ f(\boldsymbol{x})=x_1^2-2x_1-x_2,$

$$\text{s.t.}\quad 2x_1+3x_2\leqslant 6,$$
$$2x_1+x_2\leqslant 4,$$
$$x_1,\ x_2\geqslant 0.$$

解 已知数据如下

$$G=\begin{pmatrix}2&0\\0&0\end{pmatrix},\quad c=\begin{pmatrix}-2\\-1\end{pmatrix},\quad A=\begin{pmatrix}2&3\\2&1\end{pmatrix},\quad b=\begin{pmatrix}6\\4\end{pmatrix}.$$

形成线性互补问题（LCP），然后用Lemke互补转轴算法求解(详见表5-2）. 经过四次迭代，得互补基本可行解：$w=(0,10/9,0,0)^T$，$z=(1/3,0,2/3,14/9)^T$. 因为G是半正定矩

表 5-2

基	w_1	w_2	w_3	w_4	z_1	z_2	z_3	z_4	z_0	q
w_1	1	0	0	0	0	0	2	3	−1	6
w_2	0	1	0	0	0	0	2	1	−1	4
w_3	0	0	1	0	−2	−2	−2	0	(−1)	−2
w_4	0	0	0	1	−3	−1	0	0	−1	−1
w_1	1	0	−1	0	2	2	4	3	0	8
w_2	0	1	−1	0	2	2	4	1	0	6
z_0	0	0	−1	0	2	2	2	0	1	2
w_4	0	0	−1	1	−1	1	(2)	0	0	1
w_1	1	0	1	−2	4	0	0	(3)	0	6
w_2	0	1	1	−2	4	0	0	1	0	4
z_0	0	0	0	−1	3	1	0	0	1	1
z_3	0	0	−1/2	1/2	−1/2	1/2	1	0	0	1/2
z_4	1/3	0	1/3	−2/3	4/3	0	0	1	0	2
w_2	−1/3	1	2/3	−4/3	8/3	0	0	0	0	2
z_0	0	0	0	−1	(3)	1	0	0	1	1
z_3	0	0	−1/2	1/2	−1/2	1/2	1	0	0	1/2
z_4	1/3	0	1/3	−2/9	0	−4/9	0	1	−4/9	14/9
w_2	−1/3	1	2/3	−4/9	0	−8/9	0	0	−8/9	10/9
z_1	0	0	0	−1/3	1	1/3	0	0	1/3	1/3
z_3	0	0	−1/2	1/3	0	2/3	1	0	1/6	2/3

阵，所以，K－T点$x^*=(2/3, 14/9)^T$是凸二次规划的最优解．经计算得最优值$f^*=-22/9=-2.4444$.

例3 $\min\ f(x)=x_1^2-2x_1x_2+x_2^2-2x_1-4x_2$,

s.t. $-x_1+x_2\leqslant 1/4$,

$x_1-2x_2\leqslant 4$,

$x_1,\ x_2\geqslant 0$.

解 已知数据如下：

$$G=\begin{pmatrix}2 & -2\\ -2 & 2\end{pmatrix},\quad c=\begin{pmatrix}-2\\ -4\end{pmatrix},\quad A=\begin{pmatrix}-1 & 1\\ 1 & -2\end{pmatrix},\quad b=\begin{pmatrix}1/4\\ 4\end{pmatrix}.$$

形成线性互补问题，然后用Lemke互补转轴算法解之（详见表5－3）．经过两次迭代后，下一次z_3应当进基，可是与z_3对应的一列中无正元素，迭代终止．此时得一准互补基本可行解

$$(w, z, z_0)=(11/4, 8, 0, 0, 0, 0, 0, 1/2, 3).$$

表 5-3

基	w_1	w_2	w_3	w_4	z_1	z_2	z_3	z_4	z_0	q
w_1	1	0	0	0	0	0	－1	1	－1	1/4
w_2	0	1	0	0	0	0	1	－2	－1	4
w_3	0	0	1	0	1	－1	－2	2	－1	－2
w_4	0	0	0	1	－1	2	2	－2	(－1)	－4
w_1	1	0	0	－1	1	－2	－3	3	0	17/4
w_2	0	1	0	－1	1	－2	－1	0	0	8
w_3	0	0	1	－1	2	－3	－4	(4)	0	2
z_0	0	0	0	－1	1	－2	－2	2	1	4
w_1	1	0	－3/4	－1/4	－1/2	1/4	0	0	0	11/4
w_2	0	1	0	－1	1	－2	－1	0	0	8
z_4	0	0	1/4	－1/4	1/2	－3/4	－1	1	0	1/2
z_0	0	0	－1/2	－1/2	0	－1/2	0	0	1	3

按5.3.1计算步骤后面的说明确定极方向

$$d=(0,1,0,0,0,0,1,10)^T.$$

容易验证：射线

$$\tilde{R}=\{(w,\ z,\ z_0)=(11/4,8,0,0,0,0,0,1/2,3)+\lambda(0,1,0,0,0,0,1,1,0)\mid\lambda\geqslant 0\}$$

上的点均为准互补基本可行解．显然，二次规划的可行域非空，因而，本例属于可行域无界，目标函数$f(x)$可以趋于$-\infty$的情形．$(0,\ 0)$是一个可行点，显然，射线$\hat{R}=\{(0,0)+\lambda(1,1)\mid\lambda\geqslant 0\}$上的点均为可行点，容易算出沿此射线$f(x)=-6\lambda\to-\infty\ (\lambda\to+\infty)$．

例 4　试用Lemke互补转轴算法求下列二次规划的一个K－T点．

$$\begin{aligned}\min\quad & f(x)=x_1^2+3x_1x_2+2x_2^2+2x_1-6x_2,\\ \text{s.t.}\quad & 2x_1-3x_2\leqslant 8,\\ & -x_1+2x_2\leqslant 6,\\ & x_1,\quad x_2\geqslant 0.\end{aligned}$$

解　已知数据如下

$$G=\begin{pmatrix}2 & 3\\ 3 & 4\end{pmatrix},\quad c=\begin{pmatrix}2\\ -6\end{pmatrix},\quad A=\begin{pmatrix}2 & -3\\ -1 & 2\end{pmatrix},\quad b=\begin{pmatrix}8\\ 6\end{pmatrix}.$$

用Lemke互补转轴算法解之，经过两次迭代得一互补基本可行解：$w=(12.5,\ 3,\ 6.5,\ 0)^T$，$z=(0,0,0,1.5)^T$．$x^{(1)}=(0,1.5)^T$是二次规划的一个K－T点．（由于G不是半正定矩阵，所以不能断定这个点$x^{(1)}$是否为二次规划的最优解．）在点$x^{(1)}$目标函数值$f(x^{(1)})=-4.5$．

§5.4 几类特殊的凸二次规划

本节研究几类特殊的凸二次规划，它们的约束条件与式(5.2-2)～(5.2-3)不完全相同，然而经过适当变换，也可以用Lemke互补转轴算法求解.

5.4.1 等式约束问题

容易证明：m个等式约束

$$a_{i1}x_1+a_{i2}x_2+\cdots+a_{in}x_n=b_i,\quad i=1,\ \cdots,\ m$$

等价于以下$m+1$个不等式约束

$$a_{i1}x_1+a_{i2}x_2+\cdots+a_{in}x_n\leqslant b_i,\quad i=1,\ \cdots,\ m$$

$$-\left(\sum_{i=1}^{m}a_{i1}\right)x_1-\cdots\cdots-\left(\sum_{i=1}^{m}a_{in}\right)x_n\leqslant-\sum_{i=1}^{m}b_i.$$

设 μ_j ($j=1,\cdots,m+1$) 为以上不等式约束的Lagrange乘子，则$\lambda_j=\mu_j-\mu_{m+1}$ ($j=1,\cdots,m$) 为等式约束的 Lagrange乘子.

5.4.2 变量具有非零下界

设有凸二次规划

$$(\text{P}_4)\qquad \min\quad f(\boldsymbol{x})=\frac{1}{2}\boldsymbol{x}^TG\boldsymbol{x}+\boldsymbol{c}^T\boldsymbol{x},$$

$$\text{s.t.}\quad A\boldsymbol{x}\leqslant\boldsymbol{b},\ \boldsymbol{x}\geqslant\boldsymbol{r}\neq\boldsymbol{0}.$$

易知，通过平移$\boldsymbol{x}=\boldsymbol{t}+\boldsymbol{r}$，可将上述凸二次规划化为标准形式

$$\min\quad F(\boldsymbol{t})=\frac{1}{2}\boldsymbol{t}^TG\boldsymbol{t}+\boldsymbol{d}^T\boldsymbol{t},$$

$$\text{s.t.}\quad A\boldsymbol{t}\leqslant\boldsymbol{h},\ \boldsymbol{t}\geqslant\boldsymbol{0},$$

其中 $\boldsymbol{d}=G\boldsymbol{r}+\boldsymbol{c}$，$\boldsymbol{h}=\boldsymbol{b}-A\boldsymbol{r}$，$F(\boldsymbol{t})=f(\boldsymbol{x})-f(\boldsymbol{r})$.

5.4.3 只具有非负约束的情形

设凸二次规划只具有非负性约束，那么，K－T条件（5.2-4）就简化为

$$\begin{aligned}&\boldsymbol{v}-G\boldsymbol{x}=\boldsymbol{c}, &(5.4\text{-}1)\\&\boldsymbol{v}^T\boldsymbol{x}=0,\\&\boldsymbol{v},\ \boldsymbol{x}\geqslant 0,\end{aligned}$$

其中$\boldsymbol{v}$为Lagrange乘子向量．求向量$\boldsymbol{v}$与$\boldsymbol{x}$满足式（5.4-1），实际上就是线性互补问题（LCP）．此时，$M=G$，$\boldsymbol{q}=\boldsymbol{c}$，$\boldsymbol{w}=\boldsymbol{v}$，$\boldsymbol{z}=\boldsymbol{x}$．因而，可以用Lemke互补转轴算法求解．为了保证能够得到互补基本可行解，还需证明以下定理．

定理 5.4.1 设M为对称半正定矩阵．若$\boldsymbol{z}^TM\boldsymbol{z}=0$，则$(M+M^T)\boldsymbol{z}=0$，即条件（5.3-2）成立．

证 设$\boldsymbol{d}=M\boldsymbol{z}$，$\boldsymbol{z}^TM\boldsymbol{z}=0$，$\lambda>0$．因为$M$为对称半正定矩阵，所以

$$\begin{aligned}0&\leqslant(\boldsymbol{z}-\lambda\boldsymbol{d})^TM(\boldsymbol{z}-\lambda\boldsymbol{d})\\&=\boldsymbol{z}^TM\boldsymbol{z}+\lambda^2\boldsymbol{d}^TM\boldsymbol{d}-2\lambda\|\boldsymbol{d}\|^2\\&=\lambda^2\boldsymbol{d}^TM\boldsymbol{d}-2\lambda\|\boldsymbol{d}\|^2.\end{aligned}$$

在上式两端分别除以λ，然后令$\lambda\to 0^+$得$-2\|\boldsymbol{d}\|^2\geqslant 0$，因而，$\boldsymbol{d}=0$，即$M\boldsymbol{z}=0$．于是

$$(M+M^T)\boldsymbol{z}=2M\boldsymbol{z}=0.$$

假设仅含非负约束$\boldsymbol{x}\geqslant 0$的凸二次规划的最优解存在，而且每一个准互补基本可行解均非退化，则Lemke算法终止于一个互补基本可行解，从而得到凸二次规划的最优解．

5.4.4 仅含约束"$Ax\leqslant b$"的情形

设凸二次规划仅具有约束"$A\boldsymbol{x}\leqslant b$"，而不带约束"$\boldsymbol{x}\geqslant 0$"，而且$G$为对称正定矩阵．在这种情形下K－T条件为

$$\begin{aligned}\boldsymbol{y}\qquad\quad +A\boldsymbol{x}&=b, &(5.4\text{-}2)\\-A^T\boldsymbol{u}-G\boldsymbol{x}&=\boldsymbol{c}, &(5.4\text{-}3)\end{aligned}$$

$$\boldsymbol{u}^T\boldsymbol{y}=0,\ \boldsymbol{u},\ \boldsymbol{y}\geqslant 0. \qquad (5.4\text{-}4)$$

其中$\boldsymbol{y}$为松弛变量，$\boldsymbol{u}$为Lagrange乘子向量. 因为G为对称正定矩阵，所以可由式（5.4-3）将$\boldsymbol{x}$解出

$$\boldsymbol{x}=-G^{-1}(A^T\boldsymbol{u}+\boldsymbol{c}), \qquad (5.4\text{-}5)$$

代入式(5.4-2)消去$\boldsymbol{x}$.再令

$$M=AG^{-1}A^T,\ \boldsymbol{q}=\boldsymbol{b}+AG^{-1}\boldsymbol{c}, \qquad (5.4\text{-}6)$$

$$\boldsymbol{w}=\boldsymbol{y},\ \boldsymbol{z}=\boldsymbol{u}. \qquad (5.4\text{-}7)$$

K－T条件（式(5.4-2)～(5.4-4)）就可以化为线性互补问题

$$\boldsymbol{w}-M\boldsymbol{z}=\boldsymbol{q},\ \boldsymbol{w}^T\boldsymbol{z}=0,\ \boldsymbol{w},\ \boldsymbol{z}\geqslant 0.$$

因而可以用Lemke互补转轴算法求解. 求得互补基本可行解（$\boldsymbol{w}$, , $\boldsymbol{z}$），即（$\boldsymbol{y}$, $\boldsymbol{u}$），以后，再用式（5.4-5）求凸二次规划的最优解$\boldsymbol{x}^*$.

若按通常习惯，令$x_j=x_j'-x_j''$, x_j', $x_j''\geqslant 0$，将问题中的变量化为非负变量，再转化为线性互补问题求解，则问题的规模将扩大，因而，在上机过程中存储量与计算量都将增加很多.

5.4.5 举例

例 1 $\min\ f(\boldsymbol{x})=\frac{1}{2}\boldsymbol{x}^TG\boldsymbol{x}+\boldsymbol{c}^T\boldsymbol{x},$

s.t. $\boldsymbol{x}\geqslant 0.$

其中

$$G=\begin{pmatrix} 6 & 0 & -3 & -3 \\ 0 & 2 & 1 & -1 \\ -3 & 1 & 2 & 1 \\ -3 & -1 & 1 & 2 \end{pmatrix},\quad \boldsymbol{c}=\begin{pmatrix} -18 \\ 36 \\ 36 \\ 18 \end{pmatrix}.$$

解 令$M=G$, $\boldsymbol{q}=\boldsymbol{c}$, $\boldsymbol{z}=\boldsymbol{x}$. 用Lemke互补转轴算法求解. 迭代两次得互补基本可行解：$\boldsymbol{w}=(0,36,27,9)^T$, $\boldsymbol{z}=(3,0,0,0)^T$. 因而得二次规划的最优解$\boldsymbol{x}^*=(3,0,0,0)^T$.

最优值$f^*=-27$.

例 2 min $f(x)=\frac{1}{2}\boldsymbol{x}^T G\boldsymbol{x}$,

$$
\begin{aligned}
\text{s.t.}\quad & 2x_1+6x_2 \qquad -2x_4\leqslant 5,\\
& -3x_1-4x_2-x_3+2x_4\leqslant -8,\\
& x_1-2x_2+x_3+x_4\leqslant 3,
\end{aligned}
$$

其中

$$
\boldsymbol{G}=\begin{pmatrix} 2 & 1 & 0 & 0\\ 1 & 1 & 0 & 0\\ 0 & 0 & 2 & 0\\ 0 & 0 & 0 & 4\end{pmatrix}.
$$

解 首先求得

$$
\boldsymbol{G}^{-1}=\begin{pmatrix} 1 & -1 & 0 & 0\\ -1 & 2 & 0 & 0\\ 0 & 0 & 1/2 & 0\\ 0 & 0 & 0 & 1/4\end{pmatrix}.
$$

然后按照式(5.4-6)计算矩阵M与向量$\boldsymbol{q}$

$$
\boldsymbol{M}=\begin{pmatrix} 2 & 6 & 0 & -2\\ -3 & -4 & -1 & 2\\ 1 & -2 & 1 & 1\end{pmatrix}\begin{pmatrix} 1 & -1 & 0 & 0\\ -1 & 2 & 0 & 0\\ 0 & 0 & 1/2 & 0\\ 0 & 0 & 0 & 1/4\end{pmatrix}\begin{pmatrix} 2 & -3 & 1\\ 6 & -4 & -2\\ 0 & -1 & 1\\ -2 & 2 & 1\end{pmatrix}
$$

$$
=\begin{pmatrix} 53 & -29 & -49/2\\ -29 & 37/2 & 11\\ -49/2 & 11 & 55/4\end{pmatrix}
$$

$q=(5,-8,3)^T$.

再用Lemke互补转轴算法解对应的线性互补问题（见表5-4），得一互补基本可行解：$w=(0,0,1/2)^T$，$z=(1,2,0)^T$。令$u=z$，最后用式（5.4-5）算得二次规划的最优解

表 5-4

基	w_1	w_2	w_3	z_1	z_2	z_3	z_0	q
w_1	1	0	0	-53	29	49/2	-1	5
w_2	0	1	0	29	-37/2	-11	(-1)	-8
w_3	0	0	1	49/2	-11	-55/4	-1	3
w_1	1	-1	0	-82	(95/2)	71/2	0	13
z_0	0	-1	0	-29	37/2	11	1	8
w_3	0	-1	1	-9/2	15/2	-11/4	0	11
z_2	2/95	-2/95	0	-164/95	1	71/95	0	26/95
z_0	-37/95	-58/95	0	(279/95)	0	-537/190	1	279/95
w_3	-3/19	-16/19	1	321/38	0	-635/76	0	170/19
z_2	-58/279	-106/279	0	0	1	-255/279	164/279	2
z_1	-37/279	-58/279	0	1	0	-179/186	95/279	1
w_3	179/186	85/93	1	0	0	-7/31	-535/186	1/2

$$x^*=-\begin{pmatrix}1 & -1 & 0 & 0\\ -1 & 2 & 0 & 0\\ 0 & 0 & 1/2 & 0\\ 0 & 0 & 0 & 1/4\end{pmatrix}\begin{pmatrix}2 & -3 & 1\\ 6 & -4 & -2\\ 0 & -1 & 1\\ -2 & 2 & 1\end{pmatrix}\begin{pmatrix}1\\ 2\\ 0\end{pmatrix}$$

$$=(2,\ 0,\ 1,\ -1/2)^T.$$

例 3 $\min \quad f(x)=\frac{1}{2}x_1^2+\frac{1}{2}x_2^2+3x_1+2x_2,$

$$\text{s.t.} \quad x_1+x_2=4, \quad x=(x_1, x_2)^T \geqslant 0.$$

解 首先用两个约束：$x_1+x_2\leqslant 4$，$-x_1-x_2\leqslant -4$ 替换其中的等式约束，然后用Lemke互补转轴算法解之. 经过三次迭代得表 5-5，下一次迭代z_3应当进基. 由于$6/4=3/2<6/1$，所以将w_1或z_0从基中换出均可以. 若将z_0从基中换出，则得到互补基本可行解：$w=(0,0,0,0)^T$，$z=(0,4.5,1.5,2.5)^T$. 迭代终止. 最优解为$x^*=(1.5,2.5)^T$，最优值$f^*=13.75$. 若将w_1从基中换出，则将遇到“射线终止”，z_0仍然在基中. 然而，本例既不属于可行域为空集又不属于可行域无界的情形. 实际上，虽然z_0是基变量，但其值为0. 所以此时所得的准互补基本可行解实际上就是互补基本可行解. 这属于退化情形. 限于篇幅，关于退化情形的进一步探讨从略，请读者考虑.

表 5-5

基	w_1	w_2	w_3	w_4	z_1	z_2	z_3	z_4	z_0	q
w_1	1	-1	-2	2	0	0	4	0	0	6
z_0	0	-1	-1	1	0	0	2	0	1	3
z_4	0	0	1	-1	0	0	-1	1	0	1
z_2	0	-1	0	1	-1	1	1	0	0	6

§5.5 两类特殊的线性互补问题

在本节中研究两类特殊的线性互补问题，这两类问题，虽然向量q都有负分量，但不必引入人工变量z_0，可以通过简单公式算得一个互补基本可行解.

定理 5.5.1 设矩阵M的第k列无负元素，即$m_{ik}\geqslant 0$若对应于

q的负分量（$q_i<0$）都有$m_{ik}>0$，而且

$$\max\left\{\frac{-q_i}{m_{ik}}\middle| q_i<0\right\}=\frac{-q_k}{m_{kk}}, \tag{5.5-1}$$

则线性互补问题（LCP）有一个互补基本可行解

$$z_k=-q_k/m_{kk} \tag{5.5-2}$$

$$w_i=m_{ik}\cdot z_k+q_i,\quad i=1,\ \cdots,\ p,\ i\neq k, \tag{5.5-3}$$

其非基变量均为零。

证 用Lemke互补转轴算法解线性互补问题（LCP）时所列初始表如表5－6。以$-m_{kk}$为主元迭代一次后，基变量的值即为式（5.5-2）与（5.5-3）。由式（5.5-1）～（5.5-3）可知：$z_k>0$，$w_i\geqslant 0$，$i=1, 2, \cdots, p$，$i\neq k$。因而，由式（5.5-2）与（5.5-3）给出的是（LCP）的一个互补基本可行解（非基变量均为零）。

表 5-6

基	w_1	…	w_k	…	w_p	z_1	…	z_k	…	z_p	q
w_1	1	…	0	…	0	$-m_{11}$	…	$-m_{1k}$	…	$-m_{1p}$	q_1
…		…		…			…		…		…
w_k	0	…	1	…	0	$-m_{k1}$	…	$-m_{kk}$	…	$-m_{kp}$	q_k
…		…		…			…		…		…
w_p	0	…	0	…	1	$-m_{p1}$	…	$-m_{pk}$	…	$-m_{pp}$	q_p

定理 5.5.2 假设：1° 矩阵M的第k列对应于$q_i<0$的元素m_{ik}均为正，对应于$q_i=0$的元素m_{ik}均非负；

2° $t_1\leqslant t_2$，t_1，t_2由下式确定：

$$t_1=\max\left\{\frac{-q_i}{m_{ik}}\middle| q_i<0\right\},$$

$$t_2=\min\left\{\frac{q_i}{-m_{ik}}\middle| m_{ik}<0\right\}; \tag{5.5-4}$$

3° t_1（或t_2）$=-q_l/m_{lk}$.

则线性互补问题（LCP）有一个互补基本可行解：式（5.5-2），（5.5-3）[24].

（证略）

例 $\min \ f(\boldsymbol{x})=3x_1^2+2x_1x_2-\frac{1}{2}x_2^2-8x_1-5x_2$,

$$
\begin{aligned}
\text{s.t.} \quad & x_1+2x_2\leqslant 10,\\
& 2x_1-4x_2\leqslant -1,\\
& x_1,\quad x_2\geqslant 0.
\end{aligned}
$$

解 本例转化为线性互补问题时矩阵M与向量$\boldsymbol{q}$如下：

$$
M=\begin{pmatrix} 0 & 0 & -1 & -2\\ 0 & 0 & -2 & 4\\ 1 & 2 & 6 & 2\\ 2 & -4 & 2 & 1 \end{pmatrix},\qquad \boldsymbol{q}=\begin{pmatrix} 10\\ -1\\ -8\\ -5 \end{pmatrix}.
$$

矩阵M的第四列对应于$q_i<0$的元素均为正. 按式（5.5-4）计算t_1，t_2的值如下：

$$t_1=\max\{/,\ 1/4,\ 8/2,\ 5/1\}=5=-q_4/m_{44},$$

$$t_2=\min\{10/2,\ /,\ /,\ /\}=5.$$

$t_1\leqslant t_2$. 根据定理5.5.2线性互补问题（LCP）有一个互补基本可行解：

$z_4=5$，$w_1=-2\cdot 5+10=0$，$w_2=4\cdot 5-1=19$,

$w_3=2\cdot 5-8=2$，$w_4=z_1=z_2=z_3=0$.

二次规划的最优解为$\boldsymbol{x}^*=(0,\ 5)^T$. 将$\boldsymbol{x}^*$代入目标函数，算得最优值$f^*=-12.5$.

§5.6 灵敏度分析

5.6.1 线性互补问题的灵敏度分析

假设已经用Lemke互补转轴算法求得线性互补问题（LCP）的互补基本可行解，互补基（由初始表中与互补基本可行解的基变量相应的列组成的$p\times p$矩阵）为B，其逆为B^{-1}（由末表中前p列组成）．末表的最后一列$\hat{\boldsymbol{q}}\geqslant 0$．互补基本可行解中基变量的值由$\hat{\boldsymbol{q}}$给出．再设$\boldsymbol{h}_j$为末表中第$j$列，$h_{ij}$为$\boldsymbol{h}_j$的第$i$个元素．

定理 5.6.1 假设M的元素不变，仅$\boldsymbol{q}$的第r个元素有增量Δq_r．再设

$$t_1=\max\{(-\hat{q}_i/h_{ir})\mid h_{ir}>0\},\tag{5.6-1}$$

$$t_2=\min\{\hat{q}_i/(-h_{ir})\mid h_{ir}<0\}.\tag{5.6-2}$$

若Δq_r满足不等式

$$t_1\leqslant\Delta q_r\leqslant t_2\tag{5.6-3}$$

则互补基B不变，新的互补基本可行解中的基变量由下式给出（基变量的顺序不变）：

$$\hat{\boldsymbol{q}}+\Delta q_r\boldsymbol{h}_r\tag{5.6-4}$$

（若$\forall i$，$h_{ir}\leqslant 0$，则$t_1=-\infty$；若$\forall i$，$h_{ir}\geqslant 0$，则$t_2=+\infty$）

证 因为用B^{-1}左乘初始表即得末表，所以当初始表中的最后一列化为

$$\boldsymbol{q}+(0,\ \cdots,\ 0,\ \Delta q_r,\ 0,\ \cdots,\ 0)^T$$

时，末表中的最后一列化为

$$B^{-1}(\boldsymbol{q}+(0,\ \cdots,\ 0,\Delta q_r,\ 0,\ \cdots,\ 0)^T)=\hat{\boldsymbol{q}}+\Delta q_r\boldsymbol{h}_r.$$

若$h_{ir}>0$，则由式(5.6-1)及式(5.6-3)得

$$\hat{q}_i+\Delta q_r\cdot h_{ir}\geqslant 0.\tag{5.6-5}$$

若$h_{ir}<0$，则由式(5.6-2)与式(5.6-3)得式(5.6-5)．因为$\hat{q}_i\geqslant 0$，所以若$h_{ir}=0$，则式(5.6-5)也成立．总之，有

$$\hat{q}+\Delta q_r \boldsymbol{h}_r \geqslant 0.$$

所以，互补基B不变，并且由式(5.6-4)确定新的互补基本可行解．

定理 5.6.2 假设M的元素不变．若向量$\boldsymbol{q}$的改变量$\Delta\boldsymbol{q}$满足不等式：$\hat{\boldsymbol{q}}+B^{-1}\Delta\boldsymbol{q}\geqslant 0$，则互补基$B$不变，互补基本可行解中的基变量为$\hat{\boldsymbol{q}}+B^{-1}\Delta\boldsymbol{q}$，非基变量为零．

证 因为矩阵M无变化，所以用B^{-1}左乘初始表后，所得末表中仅最后一列有变化．最后一列化为

$$B^{-1}(\boldsymbol{q}+\Delta\boldsymbol{q})=B^{-1}\boldsymbol{q}+B^{-1}\Delta\boldsymbol{q}=\hat{\boldsymbol{q}}+B^{-1}\Delta\boldsymbol{q}.$$

根据假设，这一列无负元素，所以矩阵B仍然为互补基，新的互补基本可行解中的基变量由上式给出．

定理 5.6.3 假设向量$\boldsymbol{q}$不变，矩阵M中与互补基本可行解$(\boldsymbol{w},\boldsymbol{z})$的基变量$z_j$对应的列的元素均不变，而仅是与非基变量$z_j$对应的列的元素有变化，则互补基、互补基本可行解均不变（证明从略）．

5.6.2 凸二次规划的灵敏度分析

非线性规划的稳定性问题，即灵敏度分析（sensitivity analysis）比较复杂，关于严格凸二次规划（G为正定矩阵）解的稳定性问题，1973年Daniel曾给出一个定性的摄动定理，在本书末的参考文献12与15中都有介绍，此处从略．

我们知道，用Lemke互补转轴算法求解凸二次规划时是转化为线性互补问题（LCP）进行的，线性互补问题与二次规划的已知数据间的关系如式(5.2-5)所示．所以上段中关于线性互补问题的灵敏度分析可推广到凸二次规划．现沿用上段开头所作的假设与有关记号，再设二次规划的最优解为$\boldsymbol{x}^*$，而基变量$x_j^*=\hat{q}_s$，非基变量均为零．

若仅b_r有增量Δb_r，而其它系数均不变，那么根据定理5.6.1，当Δb_r满足不等式

$$t_1 \leqslant \Delta b_r \leqslant t_2$$

（其中t_1，t_2分别由式(5.6-1)与(5.6-2)确定）时，互补基不变，基变量

$$x_j^* = \widehat{q}_i + \Delta b_r \cdot h_{i\,r},$$

非基变量均为零。

若仅c_r有增量Δc_r，而其它系数均不变，那么根据定理5.6.1，当Δc_r满足不等式

$$\max\left\{\frac{-\widehat{q}_i}{h_{i,\ r+m}}\ \middle|_{h_{i,\ r+m}>0}\right\} \leqslant \Delta c_r$$

$$\leqslant \min\left\{\frac{\widehat{q}_i}{-h_{i,\ r+m}}\ \middle|_{h_{i,r+m}<0}\right\}$$

时，互补基不变，基变量

$$x_j^* = \widehat{q}_i + \Delta c_r \cdot h_{i,\,r+m}.$$

若$\boldsymbol{b}$、$\boldsymbol{c}$的分量中发生变化的分量多于一个，设$\boldsymbol{b}$、$\boldsymbol{c}$的改变量各为$\Delta\boldsymbol{b}$、$\Delta\boldsymbol{c}$，令$\Delta\boldsymbol{q} = (\Delta\boldsymbol{b},\ \Delta\boldsymbol{c})^T$，然后运用定理5.6.2.

作为定理5.6.3在凸二次规划方面的应用，给出以下定理.

定理 5.6.4 假设用Lemke互补转轴算法已经求得凸二次规划(QP)的最优解$\boldsymbol{x}^*$，互补基为B，互补基本可行解为($\boldsymbol{w}$，$\boldsymbol{z}$). 再设$\boldsymbol{b}$、$\boldsymbol{c}$无变化. 若

1° 矩阵G的元素g_{rk}($g_{kr} = g_{rk}$)有变化，但互补基B中不含M的第$m+k$与第$m+r$列（即z_{m+k}与z_{m+r}均为非基变量），而且G仍然为对称半正定矩阵；或

2° 矩阵A的元素a_{rk}有变化，但互补基B中不含M的第r列与第$m+k$列（即z_r与z_{m+k}均不是基变量），则矩阵B仍然为互补基，($\boldsymbol{w}$，$\boldsymbol{z}$)仍然为互补基本可行解，$\boldsymbol{x}^*$仍然为凸二次规划的最优

解，最优值f^*不变（若G化为非半正定矩阵，则x^*只是二次规划的一个K－T点）.（证明从略）

例 1 在§5.3中曾经用Lemke互补转轴算法求得几个凸二次规划的最优解. 关于§5.3中的例1矩阵

$$B=\begin{pmatrix}1 & 2 & 3 & 0\\ 0 & 1 & 4 & 0\\ 0 & -1 & 0 & -1\\ 0 & 0 & -1 & -4\end{pmatrix}$$

为互补基. 问当Δb_2在哪个范围内变化时，矩阵B仍然为互补基？

解 $t_1=\max\left\{-\frac{13}{17}\Big/\frac{1}{17},\ \frac{-18}{17}\Big/\frac{4}{17}\right\}=-\frac{9}{2}$,

$$t_2=\min\left\{\frac{22}{17}\Big/\frac{14}{17},\ \frac{4}{17}\Big/\frac{1}{17}\right\}=\frac{11}{7}.$$

由定理5.6.1知，当$-\frac{9}{2}\leqslant\Delta b_2\leqslant\frac{11}{7}$（即$\frac{1}{2}\leqslant b_2\leqslant\frac{47}{7}$）时，矩阵$B$仍然为互补基.

例 2 $\min\ f(x)=\frac{1}{2}x_1^2+\frac{1}{2}x_2^2-2x_1-4x_2$,

$$\begin{aligned}\text{s.t.}\quad & 2x_1+3x_2\leqslant 14,\\ & x_1+4x_2\leqslant 14,\\ & x_1,\ x_2\geqslant 0.\end{aligned}$$

解 将它与§5.3例1比较，发现仅约束的右端与目标函数的一次项系数有改变，其它系数均未变. 化为线性互补问题（LCP）时仅向量q有变化，$\Delta q=(8,9,-1,-2)^T$. 根据表5-1得

$$\hat{q}+B^{-1}\Delta q$$

$$=\begin{pmatrix}22/17\\13/17\\18/17\\4/17\end{pmatrix}+\begin{pmatrix}1&-14/17&20/17&-5/17\\0&1/17&-16/17&4/17\\0&4/17&4/17&-1/17\\0&-1/17&-1/17&-4/17\end{pmatrix}\begin{pmatrix}8\\9\\-1\\-2\end{pmatrix}$$

$$=(22/17,\ 30/17,\ 52/17,\ 4/17)^T\geqslant\mathbf{0},$$

根据定理5.6.2，互补基B不变，互补基本可行解为

$$w=(22/17,\ 0,\ 0,\ 0)^T,$$

$$z=(0,\ 4/17,\ 30/17,\ 52/17)^T.$$

二次规划的最优解为：$x^*=(30/17,\ 52/17)^T$. 经计算最优值$f^*=-9.5294$.

例3 关于§5.3中例4，已经求得互补基本可行解：$w=(12.5,\ 3,\ 6.5,\ 0)^T$，$z=(0,\ 0,\ 0,\ 1.5)^T$与K-T点$x^*=(0,\ 1.5)^T$. 由于z_1，z_2，z_3都不是基变量，所以a_{11}，a_{21}与g_{11}的变化（$g_{11}>0$）不会引起互补基本可行解的变化，$x^*=(0,\ 1.5)^T$仍然为二次规划的K-T点. 经验证，当$a_{11}=8$，$a_{21}=9$，$g_{11}=6$时，互补基本可行解的确没有改变，$x^*=(0,\ 1.5)^T$仍然是新的（严格凸）二次规划的最优解.

§5.7 凸二次规划的对偶理论

凸二次规划的对偶理论（duality theory）是凸二次规划的重要内容之一. 不同文献中所讨论的对偶规划的形式、对偶规划的理论不完全相同. 现列举对偶规划的四种形式，而以下将要讨论的对

偶定理与线性规划的对偶定理比较接近，容易掌握。

当凸二次规划的约束具有不同形式时，其对偶规划的约束也具有不同形式，其对应关系可列举如下[20]：

原始规划（QP）	对偶规划(DP)
min $f(\boldsymbol{x})=\frac{1}{2}\boldsymbol{x}^T G\boldsymbol{x}+\boldsymbol{c}^T\boldsymbol{x}$，	max $g(\boldsymbol{u},\boldsymbol{v})=-\frac{1}{2}\boldsymbol{u}^T G\boldsymbol{u}+\boldsymbol{b}^T\boldsymbol{v}$，
s.t. $A\boldsymbol{x}\geqslant\boldsymbol{b}$，	s.t. $A^T\boldsymbol{v}-G\boldsymbol{u}=\boldsymbol{c}$，$\boldsymbol{v}\geqslant 0$，
或 $A\boldsymbol{x}\geqslant\boldsymbol{b}$，$\boldsymbol{x}\geqslant 0$，	或 $A^T\boldsymbol{v}-G\boldsymbol{u}\leqslant\boldsymbol{c}$，$\boldsymbol{v}\geqslant 0$，
或 $A\boldsymbol{x}=\boldsymbol{b}$，	或 $A^T\boldsymbol{v}-G\boldsymbol{u}=\boldsymbol{c}$，
或 $A\boldsymbol{x}=\boldsymbol{b}$，$\boldsymbol{x}\geqslant 0$.	或 $A^T\boldsymbol{v}-G\boldsymbol{u}\leqslant\boldsymbol{c}$.

其中G为$n\times n$对称半正定矩阵，A为$m\times n$矩阵，$\boldsymbol{c}$，$\boldsymbol{x}$，$\boldsymbol{u}\in\mathscr{R}^n$，$\boldsymbol{b}$，$\boldsymbol{v}\in\mathscr{R}^m$（后三种可以经过第一种得到）.

在凸二次规划的对偶理论中有以下两个基本定理.

定理 5.7.1 （弱对偶）设$\boldsymbol{x}$为(QP)的可行解，$(\boldsymbol{u},\boldsymbol{v})$为(DP)的可行解，则$f(\boldsymbol{x})\geqslant g(\boldsymbol{u},\boldsymbol{v})$.

（注：本定理是定理4.4.2的特例）

定理 5.7.2 （对偶定理）

1° 设(QP)具有最优解$\boldsymbol{x}^*$，且$\boldsymbol{x}^*$为约束的正则点，则(DP)具有最优解（$\boldsymbol{x}^*$，λ^*），且$g(\boldsymbol{x}^*,\lambda^*)=f(\boldsymbol{x}^*)$；

2° 设G为正定矩阵，若（$\boldsymbol{x}^*$，λ^*）是（DP）的最优解，且为约束的正则点，则$\boldsymbol{x}^*$是（QP）的最优解，且

$$f(\boldsymbol{x}^*)=g(\boldsymbol{x}^*,\lambda^*).$$

证 1°（定理的第一部分是定理4.4.3的特殊情形）

2° 因为（$\boldsymbol{x}^*$，λ^*）是（DP）的最优解，且为约束的正则点，所以存在$\boldsymbol{y}^*$，$\boldsymbol{s}^*$满足K-T条件

$$G\boldsymbol{x}^*+G\boldsymbol{s}^*=0, \tag{5.7-1}$$

$$\boldsymbol{b}+A\boldsymbol{s}^*+\boldsymbol{y}^*=0, \tag{5.7-2}$$

$$A^T\lambda^* - Gx^* = c, \quad (5.7\text{-}3)$$

$$\lambda^{*T} y^* = 0, \quad (5.7\text{-}4)$$

$$\lambda^*,\ y^* \geqslant 0. \quad (5.7\text{-}5)$$

由式(5.7-1)得$s^* = -x^*$. 因而式(5.7-2)～(5.7-5)就是(QP)的K-T条件，所以x^*是(QP)的最优解. 再利用定理的第一部分得$f(x^*) = g(x^*, \lambda^*)$.

§5.8 Hildreth与D'espo算法

5.8.1 算法所适用的二次规划模型与理论基础

此算法适用于在§5.4中曾研究过的仅含约束"$Ax \leqslant b$"的严格凸二次规划

$$(\mathrm{QP}_1) \quad \min \ f(x) = \frac{1}{2} x^T G x + c^T x, \quad (5.8\text{-}1)$$

$$\text{s.t.} \quad Ax \leqslant b \quad (5.8\text{-}2)$$

(其中G为$n \times n$对称正定矩阵，A为$m \times n$矩阵).

由式(5.4-5)知 $x = -G^{-1}(A^T\lambda + c)$. (5.8-3)

令 $D = AG^{-1}A^T$，$q = b + AG^{-1}c$. (5.8-4)

其K-T条件就可以化为以下线性互补问题.

$$w - D\lambda = q,\ w^T\lambda = 0,\ w,\ \lambda \geqslant 0 \quad (5.8\text{-}5)$$

(其中w，$\lambda \in \mathscr{R}$分别为(5.8-2)的松弛变量与Lagrange乘子向量). 而它正是以下简单的凸二次规划

$$(\mathrm{QP}_2) \quad \min \ h(\lambda) = \frac{1}{2}\lambda^T D\lambda + q^T\lambda,\ \text{s.t.}\ \lambda \geqslant 0$$

的K-T条件. 求得(QP_2)的最优解λ^*后，将λ^*代入式(5.8-3)就可得原二次规划(QP_1)的最优解x^*.

还可以根据对偶理论分析(QP_2)与(QP_1)的关系. 凸二次规划(QP_1)(先将约束化为：$-Ax \geqslant -b$)的对偶规划是

(DP) $\max \quad g(\boldsymbol{u},\ \lambda)=-\frac{1}{2}\boldsymbol{u}^T G\boldsymbol{u}-\boldsymbol{b}^T\lambda,$ (5.8-6)

$\text{s.t.} \quad G\boldsymbol{u}+A^T\boldsymbol{\lambda}=-\boldsymbol{c},\ \boldsymbol{\lambda}\geqslant 0.$ (5.8-7)

由式(5.8-7)解出

$$\boldsymbol{u}=-G^{-1}(A^T\boldsymbol{\lambda}+\boldsymbol{c}), \tag{5.8-8}$$

代入目标函数(5.8-6)得

$$g(\boldsymbol{u},\boldsymbol{\lambda})=-\frac{1}{2}(A^T\boldsymbol{\lambda}+\boldsymbol{c})^T G^{-1}(A^T\boldsymbol{\lambda}+\boldsymbol{c})-\boldsymbol{b}^T\boldsymbol{\lambda}$$

$$=-\frac{1}{2}\boldsymbol{\lambda}^T AG^{-1}A^T\boldsymbol{\lambda}-\boldsymbol{\lambda}^T AG^{-1}\boldsymbol{c}-\boldsymbol{b}^T\boldsymbol{\lambda}-\frac{1}{2}\boldsymbol{c}^T G^{-1}\boldsymbol{c}$$

$$=-\frac{1}{2}\boldsymbol{\lambda}^T D\boldsymbol{\lambda}-\boldsymbol{q}^T\boldsymbol{\lambda}-\frac{1}{2}\boldsymbol{c}^T G^{-1}\boldsymbol{c} \quad (\text{由}5.8\text{-}4)$$

$$=-h(\boldsymbol{\lambda})-\frac{1}{2}\boldsymbol{c}^T G^{-1}\boldsymbol{c}.$$

因为$\boldsymbol{c}^T G^{-1}\boldsymbol{c}$为常数，所以$\max g(\boldsymbol{u},\ \boldsymbol{\lambda})$等价于$\min\ h(\boldsymbol{\lambda})$. ($QP_2$)实际上就是($QP_1$)的对偶规划. 求得($QP_2$)的最优解$\boldsymbol{\lambda}^*$后，将$\boldsymbol{\lambda}^*$代入式(5.8-8)得$\boldsymbol{u}^*$，($\boldsymbol{u}^*$，$\boldsymbol{\lambda}^*$)是(DP)的最优解. 根据对偶定理$\boldsymbol{u}^*$就是($QP_1$)的最优解$\boldsymbol{x}^*$. 而最优值

$$f(\boldsymbol{x}^*)=g(\boldsymbol{x}^*,\ \boldsymbol{\lambda}^*)=-h(\boldsymbol{\lambda}^*)-\frac{1}{2}\boldsymbol{c}^T G^{-1}\boldsymbol{c}. \tag{5.8-9}$$

5.8.2 计算步骤与举例

假设已将严格凸二次规划(QP_1)转化为(QP_2)，d_{ij}为矩阵D的元素($d_{ij}\neq 0$)，q_i为向量$\boldsymbol{q}$的元素. 形成以下线性方程组[4]

$$d_{i1}\lambda_1+d_{i2}\lambda_2+\cdots+d_{im}\lambda_m+q_i=0,\ i=1,\ \cdots,\ m. \tag{5.8-10}$$

1° 令 $k=0$，$\lambda_j^{(k)}=0$，$j=1,\ \cdots,\ m$.

2° $\forall i\in\{1,\ \cdots,\ m\}$， 令$\lambda_j=\lambda_j^{(k)}$，$j\neq i$，代入(5.8-10)，

解出λ_i，记为$\widehat{\lambda_i}$．取$\lambda_i^{(k+1)}=\max\{0, \widehat{\lambda_i}\}$．

3° 若$\boldsymbol{\lambda}^{(k+1)}\neq\boldsymbol{\lambda}^{(k)}$，则令$k=k+1$，转2°；否则

4° 得（QP₂）的最优解$\boldsymbol{\lambda}^*=\boldsymbol{\lambda}^{(k+1)}$．再将$\boldsymbol{\lambda}^*$代入式（5.8-8）计算$\boldsymbol{u}^*$，$\boldsymbol{u}^*$即（$QP_1$）的最优解$\boldsymbol{x}^*$．

因为$\boldsymbol{\lambda}^{(k+1)}=\boldsymbol{\lambda}^{(k)}$，所以$\boldsymbol{\lambda}^{(k)}$是线性方程组(5.8-10)的解．因而$\boldsymbol{w}=\boldsymbol{0}$，$\boldsymbol{\lambda}=\boldsymbol{\lambda}^{(k)}\geqslant\boldsymbol{0}$是线性互补问题（5.8-5）的互补基本可行解，所以$\boldsymbol{\lambda}^{(k)}$是凸二次规划（$QP_2$）的K-T点，也就是（$QP_2$）的最优解．关于算法收敛性的证明请参看书末的参考文献4．

例
$$\begin{aligned}\min\ & f(\boldsymbol{x})=x_1^2-x_1x_2+x_2^2-18x_1,\\ \text{s.t.}\ & x_1+x_2\leqslant 12,\\ & -x_1+x_2\leqslant 6,\\ & x_1,\ x_2\geqslant 0.\end{aligned}$$

解 先将约束一律化为"$\leqslant$"型，已知数据如下：

$$G=\begin{pmatrix}2 & -1\\ -1 & 2\end{pmatrix},\quad \boldsymbol{c}=\begin{pmatrix}-18\\ 0\end{pmatrix},$$

$$A=\begin{pmatrix}1 & -1 & -1 & 0\\ 1 & 1 & 0 & -1\end{pmatrix}^T,\quad \boldsymbol{b}=(12,\ 6,\ 0,\ 0)^T$$

经计算得

$$G^{-1}=\begin{pmatrix}2/3 & 1/3\\ 1/3 & 2/3\end{pmatrix},$$

$$D=AG^{-1}A^T=\begin{pmatrix}2 & 0 & -1 & -1\\ 0 & 2/3 & 1/3 & -1/3\\ -1 & 1/3 & 2/3 & 1/3\\ -1 & -1/3 & 1/3 & 2/3\end{pmatrix},$$

$$\boldsymbol{q}=\boldsymbol{b}+AG^{-1}\boldsymbol{c}=(-6,\ 12,\ 12,\ 6)^T.$$

形成线性方程组

$$\begin{cases} 2\lambda_1 - \lambda_3 - \lambda_4 - 6 = 0, \\ \frac{2}{3}\lambda_2 + \frac{1}{3}\lambda_3 - \frac{1}{3}\lambda_4 + 12 = 0, \\ -\lambda_1 + \frac{1}{3}\lambda_2 + \frac{2}{3}\lambda_3 + \frac{1}{3}\lambda_4 + 12 = 0, \\ -\lambda_1 - \frac{1}{3}\lambda_2 + \frac{1}{3}\lambda_3 + \frac{2}{3}\lambda_4 + 6 = 0, \end{cases}$$

1° 令 $k=0$, $\lambda_j^{(k)}=0$, $j=1, \cdots, 4$ (即$\boldsymbol{\lambda}^{(0)}=\mathbf{0}$).

2° $\hat{\lambda}_1=3$, $\lambda_1^{(1)}=3$; $\hat{\lambda}_2=-18$, $\lambda_2^{(1)}=0$;

$\hat{\lambda}_3=-18$, $\lambda_3^{(1)}=0$; $\hat{\lambda}_4=-9$, $\lambda_4^{(1)}=0$.

$$\boldsymbol{\lambda}^{(1)}=(3, 0, 0, 0)^T \neq \boldsymbol{\lambda}^{(0)}.$$

3° $\hat{\lambda}_1=3$, $\lambda_1^{(2)}=3$; $\hat{\lambda}_2=-18$, $\lambda_2^{(2)}=0$;

$\hat{\lambda}_3=-27/2$, $\lambda_3^{(2)}=0$; $\hat{\lambda}_4=-9/2$, $\lambda_4^{(2)}=0$.

$$\boldsymbol{\lambda}^{(2)}=(3, 0, 0, 0)^T=\boldsymbol{\lambda}^{(1)}.$$

因为 $\boldsymbol{\lambda}^{(2)}=\boldsymbol{\lambda}^{(1)}$, 所以 $\boldsymbol{\lambda}^*=\boldsymbol{\lambda}^{(2)}=(3, 0, 0, 0)^T$.

原始二次规划的最优解

$$\boldsymbol{x}^*=-G^{-1}(A^T\boldsymbol{\lambda}^*+\boldsymbol{c})$$

$$=-\begin{pmatrix} 2/3 & 1/3 \\ 1/3 & 2/3 \end{pmatrix}\begin{pmatrix} -15 \\ 3 \end{pmatrix}=\begin{pmatrix} 9 \\ 3 \end{pmatrix},$$

最优值 $f^*=9^2-9\cdot 3+3^2-18\cdot 9=-99$.

习 题

1. 取初始点$\boldsymbol{x}^{(0)}=(1, 1)^T$, 试用可行点有效集算法解§5.3中的例1.

2. 试写出与下列凸二次规划对应的线性互补问题.

1° $\min\ f(\boldsymbol{x})=2x_1^2-4x_1x_2+4x_2^2-6x_1-3x_2$,

s.t. $x_1+x_2\leqslant 3$, $4x_1+x_2\leqslant 9$, $x_1, x_2\geqslant 0$.

2° min $f(\boldsymbol{x})=x_1^2-x_1x_2+0.5x_2^2+x_1-2x_2$,

s.t. $3x_1+5x_2\leqslant 12$, x_1, $x_2\geqslant 0$.

3. 试用Lemke互补转轴算法解上题中的两个凸二次规划

4. 解下列二次规划

1° min $f(\boldsymbol{x})=x_1^2+2x_2^2+x_3^2-4x_1+2x_2$,

s.t. $2x_1+x_2+x_3=9$,

$x_1+x_2-2x_3=8$,

x_1, x_2, $x_3\geqslant 0$.

2° min $f(\boldsymbol{x})=0.5x_1^2-x_1x_2+x_2^2-x_1+x_2$,

s.t. $x_1+x_2\leqslant -2$, x_1, $x_2\geqslant 0$.

5. 试证明式(5.4-6)中的矩阵M是对称半正定矩阵，若A为列满秩矩阵，则M为正定矩阵.

6. 试证明约束为"$A\boldsymbol{x}=\boldsymbol{b}$"的凸二次规划的对偶规划可以通过约束为"$A\boldsymbol{x}\geqslant\boldsymbol{b}$"的凸二次规划的对偶规划得到.

7. 试写出习题2.1°与习题4.1°两个小题中的二次规划的对偶规划.

8. 试用Hildreth与D′espo算法解以下凸二次规划

min $f(\boldsymbol{x})=x_1^2-x_1x_2+0.5x_2^2+x_1-2x_2$,

s.t. $3x_1+2x_2\leqslant 8$, x_1, $x_2\geqslant 0$.

9. 试利用第3题的结果解以下凸二次规划

min $f(\boldsymbol{x})=2x_1^2-4x_1x_2+4x_2^2+9x_1-4x_2$,

s.t. $x_1+x_2\leqslant 5$, $4x_1+x_2\leqslant 7$, x_1, $x_2\geqslant 0$.

10. 试分析§5.4例1中G的哪些元素的改变（设$\boldsymbol{c}$保持不变）不影响互补基本可行解，也不影二次规划的最优解$\boldsymbol{x}^*$与最优值f^*？并举例验证.

11. 试编写关于Hildreth与D′espo算法的源程序.

附录1　预备知识

为方便读者，列举《数学分析》与《点集拓扑学》中的与本书有关的若干基本概念和定理，供查阅、参考.

1.　集合C的上确界与下确界

定义1　设C是上有界的实数集，则称C的最小上界μ为C的**上确界**（supremum），记为$\sup C$. $\mu=\sup C$等价于：

$1°$　$x\leqslant\mu$，$\forall x\in C$；

$2°$　对于任一正数ε均存在$x_1\in C$，使$x_1>\mu-\varepsilon$.

定义2　设C为下有界的实数集，则称C的最大下界α为C的**下确界**（infimum），记为$\inf C$. $\alpha=\inf C$等价于：

$1°$　$x\geqslant\alpha$，$\forall x\in C$；

$2°$　对于任一正数ε均存在$x_1\in C$，使$x_1<\alpha+\varepsilon$.

2.　点列$\{\boldsymbol{x}^{(k)}\}$的极限与点$\boldsymbol{x}^{(0)}$的δ邻域

定义3　设$\{\boldsymbol{x}^{(k)}\}$是$\mathscr{R}^n$中的点列，$\boldsymbol{x}\in\mathscr{R}^n$. 若$\lim\limits_{k\to\infty}\|\boldsymbol{x}^{(k)}-\boldsymbol{x}\|=0$,则称点列$\{\boldsymbol{x}^{(k)}\}$收敛于点$\boldsymbol{x}$,记为$\lim\limits_{k\to\infty}\boldsymbol{x}^{(k)}=\boldsymbol{x}$或$\boldsymbol{x}^{(k)}\to\boldsymbol{x}$（$k\to\infty$）.

定义4　以点$\boldsymbol{x}^{(0)}$为中心以$\delta>0$为半径的球$\{\boldsymbol{x}\mid\|\boldsymbol{x}-\boldsymbol{x}^{(0)}\|<\delta\}$称为点$\boldsymbol{x}^{(0)}$的$\delta$邻域，记为$U(\boldsymbol{x}^0;\ \delta)$或$B(\boldsymbol{x}^0;\ \delta)$.

3.　开集与闭集

定义5　设E为度量空间，$A\subset E$. 若$\forall\,\boldsymbol{x}\in A$均存在$r>0$使$B(\boldsymbol{x};\ r)\subset A$，则称$A$为开集（其中$r$的值依赖于$\boldsymbol{x}$）.

性质1　任意个开集的并是开集.

性质2　有限个开集的交是开集.

定义6　设$A\subset U$，而U为开集，则称U为A的开邻域. 包含A

的一个开邻域U的任意集合均称为A的邻域.

特别地，含$\boldsymbol{x}$的开集U称为点$\boldsymbol{x}$的开邻域. 包含$\boldsymbol{x}$的一个开邻域的任意集合均称为点$\boldsymbol{x}$的邻域.

定义7 设$\boldsymbol{x}\in A$. 若存在$\boldsymbol{x}$的邻域$U\subset A$，则称点$\boldsymbol{x}$为A的内点. A的全部内点称为A的内部，记为$\mathrm{int}(A)$.

定理1 A是开集$\Longleftrightarrow A=\mathrm{int}(A)$.

定义8 设E为度量空间，$A\subset E$为开集，B为A的补集：$B=E\setminus A$，则称B为闭集.

空集Φ与E本身既是开集，又是闭集.

性质3 任意个闭集的交是闭集.

性质4 有限个闭集的并是闭集.

定义9 含A的最小闭集称为A的闭包，记为$\overline{A}$.

定理2 A是闭集$\Longleftrightarrow A=\overline{A}$.

定义10 若$\boldsymbol{x}\in\overline{A}\setminus\mathrm{int}(A)$,则称$\boldsymbol{x}$为$A$的边界点. A的边界点的全体称为A的边界（boundary），记为$\mathrm{bd}(A)$.

4. 列紧集与紧集

定义11 设A是度量空间E的子集. 若A的每一序列$\{\boldsymbol{x}^k$都存在收敛于A中一点的子列，则称A为**列紧集**（sequentially compact set）.

性质5 设A为n维欧氏空间$\mathscr{R}^n$中的有界闭集，则A为列紧集.

性质6 设A、B为列紧集，则直积$A\times B$与并$A\cup B$均为列紧集.

性质7 列紧集A的非空闭子集是列紧集.

性质8 列紧集A的连续象集是列紧集.

定义12 若度量空间E本身是列紧集,则称E为列紧空间（sequentially compact space）.

n维欧氏空间$\mathscr{R}^n$不是列紧空间.

定义13 设E为度量空间．若E的覆盖由开集组成，则称它为E的开覆盖．若E的任一开覆盖$\{U_\lambda|\lambda\in L\}$中均存在$E$的有限子覆盖$\{U_\lambda|\lambda\in H\}$其中$H$为$L$的有限子集，则称$E$为紧空间（compact space）.

定理3 设E为度量空间．则E是紧空间$\Longleftrightarrow E$是列紧空间.

定义14 若度量空间E的子空间A是紧空间，则称A为**紧集**（compact set）.

注 在n维欧氏空间$\mathscr{R}^n$中所论紧集指有界闭集.

5. 中值定理与泰勒公式

定理4（中值定理） 设$D\subset\mathscr{R}^n$为开凸集，在D内$f(\boldsymbol{x})\in C^1$，$\boldsymbol{x}$，$\boldsymbol{y}\in D$，则存在$\theta\in(0,1)$使

$$f(\boldsymbol{y})=f(\boldsymbol{x})+\nabla f(\boldsymbol{\xi})^T(\boldsymbol{y}-\boldsymbol{x})$$

其中$\boldsymbol{\xi}=\theta\boldsymbol{x}+(1-\theta)\boldsymbol{y}$.

定理5（泰勒公式） 设$D\subset\mathscr{R}^n$为开凸集，在D内$f(\boldsymbol{x})\in C^2$，$\boldsymbol{x}$，$\boldsymbol{y}\in D$，则存在$\theta\in(0,1)$使

$$f(\boldsymbol{y})=f(\boldsymbol{x})+\nabla f(\boldsymbol{x})^T(\boldsymbol{y}-\boldsymbol{x})+\frac{1}{2}(\boldsymbol{y}-\boldsymbol{x})^TH(\boldsymbol{\xi})(\boldsymbol{y}-\boldsymbol{x})$$

其中 $\boldsymbol{\xi}=\theta\boldsymbol{x}+(1-\theta)\boldsymbol{y}$，$H(\boldsymbol{x})$为$f(\boldsymbol{x})$的Hesse矩阵，即

$$H(\boldsymbol{x})=\begin{pmatrix}\dfrac{\partial^2 f}{\partial x_1^2} & \dfrac{\partial^2 f}{\partial x_1\partial x_2} & \cdots\cdots & \dfrac{\partial^2 f}{\partial x_1\partial x_n}\\ \dfrac{\partial^2 f}{\partial x_2\partial x_1} & \dfrac{\partial^2 f}{\partial x_2^2} & \cdots\cdots & \dfrac{\partial^2 f}{\partial x_2\partial x_n}\\ \cdots & \cdots & \cdots & \cdots\\ \dfrac{\partial^2 f}{\partial x_n\partial x_1} & \dfrac{\partial^2 f}{\partial x_n\partial x_2} & \cdots\cdots & \dfrac{\partial^2 f}{\partial x_n^2}\end{pmatrix}.$$

注 $f(\boldsymbol{x})$的Hesse矩阵$H(\boldsymbol{x})$是对称的．有时将它记为$\nabla^2 f(\boldsymbol{x})$，$\left(\dfrac{\partial^2 f}{\partial x_i\partial x_j}\right)$或$(f_{ij})$.

附录2　主要记号索引

$\boldsymbol{C}^a$	集合C的代数闭包	15
$\mathrm{ri}(C)$	集合C的相对内部	45
$\mathrm{rb}(C)$	集合C的相对边界	45
$\mathrm{co}(A)$	集合A的凸包	4
$\overline{\mathrm{co}}(A)$	集合A的闭凸包	24
$\dim(A)$	集合A的维数	8
$\mathrm{aff}(C)$	集合C的仿射包	8
K^0	锥K的极	29
K^{00}	锥K的双极	29
P^m	$\mathscr{R}^m$中的非负卦限	30
$d(A, B)$	集合A，B间的距离	41
$d(\boldsymbol{a}, B)$	点$\boldsymbol{a}$与集合B间的距离	41
$a(A, B)$	$\sup\{d(\boldsymbol{x}, B) \mid \boldsymbol{x} \in A\}$	41
$h(A, B)$	集合A，B间的Hausdorff距离	41
$(\boldsymbol{x} \mid \boldsymbol{u})$	E上的连续线性函数$u(\boldsymbol{x})$	28
$(A \mid \boldsymbol{u}) \geqslant a$	$(\boldsymbol{x} \mid \boldsymbol{u}) \geqslant \alpha$，$\forall \boldsymbol{x} \in A$	28
$f(C) \geqslant \alpha$	$f(\boldsymbol{x}) \geqslant \alpha$，$\forall \boldsymbol{x} \in C$	52
$(\boldsymbol{x} \mid K^0) \leqslant 0$	$(\boldsymbol{x} \mid \boldsymbol{u}) \leqslant 0$，$\forall \boldsymbol{u} \in K^0$	29
T，T^{-1}	线性映射，逆映射	3
$\mathrm{dom}(f)$	凸函数$f(\boldsymbol{x})$的有效域	74
$\mathrm{epi}(f)$	$f(\boldsymbol{x})$的上图	75
$\delta_A(\boldsymbol{x})$	A的指示函数	90
$f \mid m$	$f(\boldsymbol{x})$在直线m上的限制	78
$\lim\limits_{x \to a} \inf f(\boldsymbol{x})$	$\sup\{\inf f(\boldsymbol{x}) \mid \boldsymbol{x} \in U \setminus \{\boldsymbol{a}\}\}$	89
$\boldsymbol{a}^T\boldsymbol{b}$	向量$\boldsymbol{a}$与$\boldsymbol{b}$的内积	189
$L(f, \beta)$	水平集	83
$\overline{f}$	$f(\boldsymbol{x})$的下半连续包	98

附录3　凸二次规划源程序

1. 凸二次规划的数学模型

$$\min \quad f(\boldsymbol{x})=\frac{1}{2}\boldsymbol{x}^T G\boldsymbol{x}+\boldsymbol{c}^T\boldsymbol{x}$$

$$\begin{aligned}
\text{s.t.} \quad & a_{i1}x_1+a_{i2}x_2+\cdots+a_{in}x_n\leqslant b_i,\ i=1,\ \cdots,\ p_1\\
& a_{i1}x_1+a_{i2}x_2+\cdots+a_{in}x_n\geqslant b_i,\ i=p_1+1,\cdots,p_1+p_2\\
& a_{i1}x_1+a_{i2}x_2+\cdots+a_{in}x_n=b_i,\ i=p_1+p_2+1,\\
& \qquad\qquad\qquad\qquad\qquad\qquad \cdots,\ p_1+p_2+p_2\\
& \boldsymbol{x}=(x_1,\ x_2,\ \cdots,\ x_n)^T\geqslant 0
\end{aligned}$$

其中G为$n\times n$对称半正定矩阵，$\boldsymbol{c}\in\mathscr{R}^n$，$b_i$为实数．$p_1$，$p_2$，$p_3$为非负整数．

2. 功能与适用范围

本程序根据Lemke互补转轴算法用BASIC语言写成．可 以在微机上求解凸二次规划．若有最优解，则可以输出最优解 X^*、最优值F^*与迭代次数I_0；若无最优解，则输出标志："WJ"．要求变量个数与线性约束个数总和不超过30；若超过此限，应修改1500定义语句中数组的维数．

3. 原始数据的输入

可以在100～300行将原始数据置入DATA语句．因为矩阵G具有对称性，所以主对角线左下方的元素不必输入．原始数据在DATA语句中的顺序如下（设g_{ij}为G的元素，c_j为$\boldsymbol{c}$的元素）：

100　DATA　n，p_1，p_2，p_3

110　DATA　g_{11}，g_{12}，$\cdots$，g_{1n}，c_1

```
120   DATA   g22, g23, g2n, c2
...............................
160   DATA   gnn, cn
170   DATA   a11, a12, …, a1n, b1
180   DATA   a21, a22, …, a2n, b2
```
（需要几行占用几行）

4. 举例

例 1 min $f(\boldsymbol{x})=\frac{1}{2}x_1^2+\frac{1}{2}x_2^2-x_1-2x_2$,

s.t. $2x_1+3x_2\leqslant 6$, $x_1+4x_2\leqslant 5$, x_1, $x_2\geqslant 0$.

解 原始数据已置入源程序前面的DATA语句，此处从略．运行后输出结果：最优解$X^*=(0.7647,1.0588)^T$，最优值$F^*=-2.0294$，迭代次数$I_0=4$（请与§5.3例1对照）．

例 2 min $f(\boldsymbol{x})=x_1^2-2x_1x_2+x_2^2-2x_1-4x_2$,

s.t. $-x_1+x_2\leqslant 0.25$, $x_1-2x_2\leqslant 4$, x_1, $x_2\geqslant 0$.

解 根据本题中的已知数据修改程序前面的DATA语句．运行后输出标志“WJ”．经检验本例的可行域非空，所以属于可行域无界，无有限的最优解的情形（请与§5.3例3对照）．

5. 凸二次规划源程序

```
100   DATA   2, 2, 0, 0
110   DATA   1, 0, -1
120   DATA   1, -2
130   DATA   2, 3, 6
140   DATA   1, 4, 5
.
.
1060  READ  N, P1, P2, P3
1070  IF P3=0  GOTO  1100
1080  P0=P1+P2+P3+1
```

```
1090  GOTO  1105
1100  P0=P1+P2
1105  P=N+P0
1110  K0=2*P+1
1120  DIM  G(N, N), C(N)
1130  FOR  I=1 TO  N
1190  FOR  J=I TO  N
1200  READ  G(I, J)
1210  G(J, I)=G(I, J)
1220  NEXT  J
1230  READ  C(I)
1240  NEXT  I
1242  IF P0=0  GOTO  1500
1250  DIM  A(P0, N), B(P0)
1260  FOR  I=1  TO  P1+P2+P3
1270  FOR  J=1  TO  N
1280  READ  A(I, J)
1290  NEXT  J
1300  READ  B(I)
1310  NEXT  I

1320  IF P2=0  GOTO 1390
1330  FOR  I=P1+1  TO  P1+P2
1340  FOR  J=1  TO  N
1350  A(I, J)=-A(I, J)
1360  NEXT  J
1370  B(I)=-B(I)
1380  NEXT I
```

```
1390  IF P3=0  GOTO 1500
1400  FOR  J=1  TO   N
1410  A(P0, J)=0
1420  NEXT J
1430  B(P0)=0
1440  FOR I=P1+P2+1  TO  P0-1
1450  FOR  J=1  TO   N
1460  A(P0, J)=A(P0, J)-A(I, J)
1470  NEXT  J
1480  B(P0)=B(P0)-B(I)
1490  NEXT  I

1500  DIM  U(31, 63),  Q(31)
1510  FOR  I=1   TO   P
1520  FOR  J=1  TO   P
1530  IF  J=I GOTO 1560
1540  U(I, J)=0
1550  GOTO  1570
1560  U(I,  J)=1
1570  NEXT  J
1580  NEXT  I
1590  IF  P0=0  GOTO 1710
1600  FOR  I=1  TO  P0
1610  U(I,  0)=I
1620  FOR  J=P+1  TO  P+P0
1630  U(I, J)=0
1640  NEXT  J
1650  FOR  J=P+P0+1  TO  2*P
```

```
1660  U(I, J)=A(I, J-P-P0)
1670  NEXT  J
1680  U(I, K0)=-1
1690  Q(I)=B(I)
1700  NEXT  I
1710  FOR  I=P0+1  TO   P
1720  U(I, 0)=I
1730  IF  P0=0  GOTO 1770
1740  FOR  J=P+1  TO P+P0
1750  U(I, J)=-A(J-P, I-P0)
1760  NEXT  J
1770  FOR  J=P+P0+1  TO  2*P
1780  U(I, J)=-G(I-P0, J-P-P0)
1790  NEXT  J
1800  U(I, K0)=-1
1810  Q(I)=C(I-P0)
1820  NEXT  I

1840  K=K0
1850  L=-2
1860  Q1=0
1870  FOR  I=1  TO   P
1880  IF  Q(I)>=Q1 GOTO  1910
1890  Q1=Q(I)
1900  L=I
1910  NEXT  I
1920  IF L=-2 GOTO 2950
1960  V=U(L, 0)
```

```
1970 U(L, 0)=K
1990 W0=U(L, K)
2000 FOR J=1 TO K0
2010 U(L, J)=U(L, J)/W0
2020 NEXT J
2030 Q(L)=Q(L)/W0
2040 FOR I=1 TO P
2050 IF I=L GOTO 2110
2060 W0=U(I, K)
2070 FOR J=1 TO K0
2080 U(I, J)=U(I, J)-W0*U(L, J)
2090 NEXT J
2100 Q(I)=Q(I)-W0*Q(L)
2110 NEXT I
2120 I0=I0+1

2170 IF V=K0 GOTO 2520
2180 IF V>P GOTO 2210
2190 K=V+P
2200 GOTO 2220
2210 K=V-P
2220 L=-1
2230 FOR I=1 TO P
2240 IF U(I, K)<=0.0001 GOTO 2290
2250 S1=Q(I)/U(I, K)
2260 IF L=-1 GOTO 2280
2270 IF S1 >=Q(L)/U(L, K) GOTO 2290
2280 L=I
```

```
2290  NEXT  I
2300  IF L=-1  GOTO  2520
2330  GOTO  1960

2520  DIM  W(P), Z(P+1), X(N), VC(N)
2522  FOR  J=1  TO   P
2524  W(J)=0
2525  Z(J)=0
2526  NEXT  J
2528  Z(P+1)=0
2530  FOR  I=1  TO   P
2540  IF U(I, 0)>P GOTO  2580
2550  T=U(I, 0)
2560  W(T)=Q(I)
2570  GOTO  2600
2580  T=U(I, 0)-P
2590  Z(T)=Q(I)
2600  NEXT  I
2602  IF L=-1 GOTO  2690
2630  GOTO  2800
2692  Z0 = Z ( P+1 )
2710  IF ABS (Z0)<0.00001 GOTO  2800
2720  GOTO  2980
2800  FOR  J=1  TO  N
2810  X(J)=Z(P0+J )
2820  NEXT  J
2830  F=0
2840  FOR  J=1  TO   N
```

```
2850  VC( J )= 0
2860  FOR  I = 1  TO   N
2870  VC( J )=VC( J )+X(I)*G(I, J)
2880  NEXT  I
2890  VC(J)=VC(J)/ 2 +C(J)
2900  F=F+VC(J) * X(J)
2910  NEXT  J

2920  PRINT  "X* : "
2922  FOR  J = 1  TO   N
2924  PRINT  X( J ),
2926  NEXT  J
2928  PRINT  "    #"
2930  PRINT  "F*=" ; F
2940  GOTO  2990
2950  PRINT  "W=Q, Z=0 "
2960  PRINT "X*=0 , F*=0 "
2970  GOTO  2990
2980  PRINT  " WJ"
2990  PRINT  "I0=" ; I0
3000  PRINT  "*---*---*"
3600  END
```